

**Universidad Miguel Hernández de Elche**  
**Facultad de Ciencias Sociales y Jurídicas de**  
**Elche**

**Titulación de Periodismo**

**Trabajo Fin de Grado**

**Curso Académico 2017-2018**



**Innovación en periodismo de datos: estudio de caso de**  
**FiveThirtyEight**

*Innovation in data journalism: a case study of*  
*FiveThirtyEight*

Alumno/a: Alberto José Losa Garrigós

Tutor/a: Prof. Miguel Carvajal Prieto

## ÍNDICE

<b>Resumen.....</b>	<b>2</b>
<b>Palabras Clave.....</b>	<b>2</b>
<b>Abstract.....</b>	<b>2</b>
<b>Key Words.....</b>	<b>2</b>
<b>1. Introducción.....</b>	<b>3</b>
<b>2. Estado de la Cuestión.....</b>	<b>5</b>
<b>3. Metodología.....</b>	<b>9</b>
<b>4. Resultados.....</b>	<b>12</b>
<b>4.1. Organización.....</b>	<b>12</b>
<b>4.2. Contenidos.....</b>	<b>13</b>
<b>4.3. Innovación y rasgos distintivos.....</b>	<b>18</b>
<b>4.4. Audiencia.....</b>	<b>21</b>
<b>5. Conclusiones.....</b>	<b>24</b>
<b>6. Notas.....</b>	<b>26</b>
<b>7. Bibliografía.....</b>	<b>27</b>
<b>8. Anexos.....</b>	<b>31</b>
<b>8.1. Entrevista a Ritchie King.....</b>	<b>31</b>
<b>8.2. Entrevista a Kiko Llaneras.....</b>	<b>35</b>
<b>8.3. Entrevista a John Burn-Murdoch.....</b>	<b>42</b>
<b>8.4. Entrevista a Daniele Grasso.....</b>	<b>46</b>
<b>8.5. Entrevista a Alberto Cairo.....</b>	<b>49</b>

## **Resumen**

El periodismo de datos es un fenómeno creciente en los últimos años. Las herramientas digitales como los lenguajes de programación, los *software* de visualización y las hojas de cálculo son cada vez más accesibles y permiten a los periodistas desarrollar historias con ellas. En 2016, la historia más leída en cualquier medio de comunicación a nivel mundial fue el modelo de agregación de encuestas de FiveThirtyEight.com para las elecciones presidenciales estadounidenses de ese año. El propósito de este Trabajo de Fin de Grado es analizar el contenido y las estrategias de innovación que utiliza FiveThirtyEight en el proceso de su actividad como medio de comunicación que tiene los datos como su nicho de mercado.

## **Palabras clave**

Periodismo de datos, innovación en periodismo, *FiveThirtyEight*, periodismo computacional, periodismo basado en datos.

## **Abstract**

Data journalism is a growing phenomenon in recent years. Digital tools such as programming languages, visualization software and spreadsheets are becoming more accessible and allow journalists to develop stories with them. In 2016, the most read story in the world was the survey aggregation model of FiveThirtyEight.com for the United States presidential election of that year. The purpose of this Final Degree Project is to analyze the content and the strategies used by FiveThirtyEight in the process of its activity as a mass media that has data as its market niche.

## **Key words**

Data journalism, innovation in journalism, *FiveThirtyEight*, computational journalism, data-driven journalism

## 1. Introducción

El uso de los datos en los medios de comunicación es un fenómeno creciente en los últimos años. Esto sucede por la variedad de usos que pueden tener en el contexto actual. Por un lado, los datos son una herramienta que permite a los medios desarrollar blogs y páginas web, analizar el tráfico de los mismos o ser parte de la toma de decisiones de las empresas. Por otro lado, los datos son un instrumento periodístico que puede tener diferentes aplicaciones como ser una fuente para desarrollar una historia o ser utilizados en infografías que permiten al usuario interactuar con la información (Royal & Blasingame, 2015).

En 2008, Nate Silver irrumpió en el panorama del periodismo de datos al desarrollar un nuevo modelo para predecir los resultados de las elecciones estadounidenses de ese año. Este modelo era una adaptación de otro que ya utilizaba en béisbol en sus análisis como *sabermétrico*<sup>1</sup> a las mediciones políticas sobre intención de voto. La diferencia respecto a sus competidores estuvo en que utilizó las encuestas de los estados más similares para completar los resultados de aquellos que no tenían una cantidad suficiente de sondeos. Además, utilizaba otros parámetros como la adición de un valor de importancia a las encuestas que se podían considerar más fiables. Esta variable fue llamada por Nate Silver como “The 538 Regression Estimate”<sup>2</sup> (Silver, 2008).

Las predicciones que realizó con este nuevo modelo resultaron ser de lo más acertadas. En su última actualización, realizada en la mañana del mismo día de las elecciones y publicada en FiveThirtyEight.com, el modelo había pronosticado entre 349 y 353 votos electorales para Barack Obama, que ganaría a John McCain. El resultado de los comicios fue una victoria de Obama con 365. Además, el modelo de Silver pronosticó el resultado exacto en 48 de los 50 estados del país norteamericano, todos menos Indiana y uno de los distritos congresuales de Nebraska.

Desde entonces, el sitio web ha pasado por distintas fases que lo han convertido en un medio de comunicación que basa sus historias en los datos, llegando a ser parte de The New York Times, ESPN y ahora de ABC News, aunque se mantiene como una redacción independiente que tiene por redactor jefe a Silver.

El propósito de este Trabajo de Fin de Grado es analizar el contenido y las estrategias que utiliza FiveThirtyEight en el proceso de su actividad como medio de comunicación

que utiliza los datos como su nicho. Sus diferentes innovaciones, su desarrollo como medio de comunicación en la última década, su estructura y los diferentes reconocimientos que ha recibido lo convierten en un caso de interés profesional y académico.

Algunas de sus aportaciones, como los diferentes modelos matemáticos para elaborar pronósticos políticos o los usos de herramientas de puntuación tradicionales, como el Elo de ajedrez, aplicados a otras disciplinas se han convertido en referencia a nivel mundial. Varios medios internacionales como El País en España utilizan técnicas basadas en estos nuevos modelos.

Las hipótesis responden a los siguientes planteamientos: FiveThirtyEight es el único medio que trata áreas tan variadas como política, deporte, cultura, ciencia y economía utilizando como fuente primaria los datos. Su crecimiento en la última década, desde su fundación en 2008, le ha convertido en un medio solvente. El hecho de que siempre ha estado ligado a una gran corporación periodística hace pensar que no sería viable como medio independiente y que su marca representa un producto de prestigio para grandes medios de comunicación.

Los objetivos de este trabajo serán analizar el contenido y las estrategias que utiliza, así como responder a la cuestión de si es viable un medio de comunicación que utiliza exclusivamente los datos como fuente y al resto de premisas anteriores. Para ello, la metodología empleada consistirá en la observación durante tres semanas de su sitio web, la evaluación del medio a través de entrevistas a expertos y la recopilación de datos contrastables como su volumen de audiencia o el tamaño de su plantilla.

## 2. Estado de la cuestión

Aunque el periodismo de datos es un fenómeno moderno, los estudios en campos relacionados llevan dándose desde la década de 1990. En 1991, Philip Meyer escribió: “Saber lo que hacer con los datos es la esencia del nuevo periodismo de precisión” (Meyer, 1991).

La expresión “periodismo de datos” es relativamente reciente, y es usada sobre todo por los medios de comunicación o blogs que hablan sobre las tendencias del periodismo, y es algo nueva como para que haya muchos artículos de investigación académica. Según Fink y Anderson (2015), es un término profundamente cuestionado y, al mismo tiempo, difuso, lo que parece poner dificultades analíticas para aquellos que quieren estudiarlo.

Muchas de las investigaciones realizadas se dedican a intentar definir y fijar un vocabulario para clasificar las formas de periodismo que tienen relación tanto con el ámbito digital como con el de la información. Algunas de las categorías en las que se enmarcan estos términos son: “periodismo de datos” (Lewis & Westlund, 2015), “periodismo computacional” (Holovaty, 2006) y “periodismo estructurado” (Chua, 2010).

Además, varios de los estudios disponibles se centran en definir qué es el periodismo de datos:

- “El periodismo de datos es la práctica de encontrar historias en los números y utilizar esos números para contar historias. (...) Es también una herramienta que integra informática, estadísticas y décadas de aprendizaje de las ciencias sociales para dar sentido a grandes bases de datos” (Howard, 2014).
- “El periodismo de datos es el resultado de noticias que involucran al usuario y que a menudo usan una base de datos para poblar la información” (Royal, 2012).
- “Lo que hace al periodismo de datos diferente es, tal vez, que permite al periodista la posibilidad de combinar el tradicional ‘olfato para las noticias’ y la capacidad de contar una historia convincente con el rango de información digital disponible” (Gray, Chambers & Bounegru, 2012).

- “Lo que hace que el periodismo de datos moderno sea diferente es la capacidad de presentar la información al mismo tiempo que da a los usuarios la posibilidad de interactuar con ella, convirtiéndola en información personalizada para cada uno” (Niles, 2006).

No todas las aplicaciones de los datos en un medio son periodismo de datos. Lo que lo diferencia de otros usos es que un volumen mayor o menor de información se acaba convirtiendo en historia (Wojcik, 2015). Normalmente, esta información no está estrictamente relacionada con el periodismo y procede de números y estadísticas y, al transformarse en un relato visual o textual, se convierte en noticia (Freixa, Pérez-Montoro & Codina, 2017).

Los productos informativos que se pueden generar como resultado de una actividad de periodismo de datos son diversos, y según varios autores incluyen términos que se pueden agrupar en “gráficos”, “infografías”, “personalización”, “mapas” y “aplicaciones web”. Distintas afirmaciones describen el periodismo de datos al hacer referencia a diferentes campos académicos. Áreas como ciencias sociales, estadística, análisis de datos, ciencia de datos e informática. Estas afirmaciones hacían énfasis en el periodismo como nexo entre los campos (Royal & Blasingame, 2015).

Para los autores Peiró y Guallar (2013), el periodismo de datos es la práctica que utiliza como materia fundamental los datos y no otras fuentes de información con tal de crear productos informativos diversos. Añaden estos autores que estos productos se pueden clasificar en tres categorías:

- Artículos basados en datos o que los contengan.
- Infografías y visualizaciones basadas en datos.
- Bases de datos y conjuntos de datos en abierto.

Estas diferencias que genera el periodismo de datos respecto al periodismo tradicional han llevado a discusiones acerca de si se puede considerar al periodismo de datos como periodismo. Adrian Holovaty planteó este debate en una entrada en su blog con cierto tono sarcástico. Ante las preguntas “¿son los datos periodismo?” y “¿es periodismo publicar una base de datos sin procesar?” respondió de la siguiente manera: “¿A quién le importa? Espero que mis competidores pierdan el mayor tiempo posible en discutir

sobre esto?” (Holovaty, 2009).

Por otra parte, en una era en la que, según las encuestas, los medios de comunicación no tienen la confianza de los ciudadanos tanto en España (CIS, 2015) como en Estados Unidos (Sullivan, 2017), el periodismo de datos aporta una visión científica a las historias, de manera que los usuarios o la audiencia pueden realizar el mismo proceso que el autor y llegar a una conclusión idéntica (Crucianelli, 2012). Pese a los aspectos positivos mencionados, los datos no son la solución a todos los problemas del periodismo, sino que abren un camino que asegura un valor añadido a la hora de elaborar los relatos informativos (Perry & Paz, 2014).

En este contexto de innovación, cabe destacar la aparición de nuevas startups enfocadas en temas. Según Carlson y Usher (2015), estas no buscan romper la autoridad del periodismo tradicional, sino romper la barrera entre periodismo y tecnología y abogar por conexiones más cercanas entre los periodistas y los ingenieros. En ránking de innovación periodística de la Universidad Miguel Hernández del año 2014, el 56% de las innovaciones en España presentaba una base tecnológica (De-Lara-González et al., 2015).

Alberto Cairo, doctor en periodismo y profesor en la Universidad de Miami, opina que la innovación tiene que orientarse a utilizar correctamente los recursos existentes, más que a desarrollar tecnología nueva: “La innovación en periodismo no debería centrarse en crear herramientas o métodos nuevos, sino en centrarse en entender lo que tenemos”.

Medios de comunicación tradicionales de gran importancia como El Mundo en España o The Guardian en Reino Unido se han convertido en referentes a nivel internacional en el trabajo con datos y su visualización (López-García, Toural-Bran & Rodríguez-Vázquez, 2016).

El cuestionario interactivo “*How Ya’ll, Youse and You Guys Talk*” fue el artículo con más visitas de The New York Times en el año 2013, y consistía en 25 preguntas sobre diferentes variedades dialectales con un plano de los Estados Unidos que, tras cada respuesta del usuario, iba delimitando con un mapa de calor la zona en la que utilizaban las expresiones más parecidas en base a los datos recopilados previamente (Wilson & Katz, 2013). El diario neoyorquino es uno de los medios tradicionales que más apuesta por el periodismo de datos. En 2014 creó la sección The Upshot, especializada en



periodismo de datos para analizar la actualidad política y económica tras la marcha de FiveThirtyEight.

Es este medio, FiveThirtyEight.com, el que más destaca por haber hecho del análisis de datos su nicho. El sitio estadounidense, que formó parte de The New York Times previo a la creación de The Upshot y ahora pertenece a ABC News, comenzó como un blog dedicado a realizar encuestas y predicciones políticas en 2008 formado por dos personas. En 2012, cuando todavía pertenecía al New York Times, FiveThirtyEight registró el 20% de todo el tráfico del diario neoyorquino durante la campaña para las elecciones presidenciales del mismo año (McDuling, 2014).

En la actualidad, el medio tiene una plantilla de más de una treintena de personas y recibe más de diez millones de usuarios únicos al mes, con entradas sobre política, deportes, economía, cultura, ciencia y salud. Todo su contenido está basado en el uso de los datos, en distintos productos como los mencionados por los autores Peiró y Guallar.

Su trabajo le ha llevado a recibir numerosos reconocimientos, siendo uno de los más recientes el premio al “Sitio Web de Periodismo de Datos del Año” en 2016 por la Global Editions Network, organización dedicada a promover la innovación en periodismo. Además, algunas de sus piezas concretas como “Swing the Election” en 2016 o “The Atlas of Redistricting” en 2018 también han recibido distinciones. Todo esto hace que FiveThirtyEight sea considerado un caso de estudio.

### 3. Metodología

La selección de este caso responde a las siguientes premisas:

- El periodismo de datos en Estados Unidos está más desarrollado que en otros países. Esto ha dado lugar a que se desarrollen diferentes iniciativas innovadoras. FiveThirtyEight es la única de todas ellas que desarrolla los principales temas de un medio generalista (actualidad, política, deporte, economía, ciencia y cultura) desde la perspectiva del periodismo de datos.
- FiveThirtyEight es el medio pionero en cuanto a la apuesta por los datos como nicho del medio de comunicación.
- Este medio también es referencia en cuanto al uso de los datos como principal fuente de sus contenidos.

Este trabajo tiene como propósito principal analizar el caso de FiveThirtyEight.com y averiguar si es viable un medio de comunicación basado en el periodismo de datos, dado que FiveThirtyEight lleva diez años online, los últimos cuatro como un medio que cubre temas más allá de la política; así distinguir los rasgos que diferencian a este medio de otros que utilizan el periodismo de datos. Desde un punto de vista más general, también se pretende conocer la situación actual del periodismo de datos y las habilidades necesarias para desarrollarlo.

Para dichos objetivos, se plantea un método de análisis basado en un estudio cualitativo del medio FiveThirtyEight. Para ello, se han analizado las estrategias y el contenido del medio. El motivo es averiguar cuáles son las directrices que sigue al estar especializado en periodismo de datos, de dónde obtiene sus fuentes y cómo componen sus temas. También se ha seguido su crecimiento en las distintas etapas de la última década, para comprobar su evolución. Por último, se han buscado sus estrategias de cara de futuro, estudiando sus innovaciones más recientes. Para esto, el trabajo cuenta con una entrevista a Ritchie King, de FiveThirtyEight, y cuatro a periodistas especializados en datos que no trabajan en el medio, a modo de fuentes externas.

King es el coordinador del área de visualización en FiveThirtyEight. Su entrevista siguió un formato similar a las demás pero se añadieron preguntas específicas al funcionamiento de FiveThirtyEight, así como sus objetivos y organización. La duración

de la misma fue de treinta minutos y fue realizada mediante videollamada.

Los periodistas especializados en datos que han sido entrevistados son Kiko Llaneras, John Burn-Murdoch, Daniele Grasso y Alberto Cairo. Estas entrevistas duraron entre treinta minutos y una hora, en función de la disponibilidad del entrevistado. Las entrevistas a Llaneras y Grasso fueron realizadas por vía telefónica, mientras que las de Burn-Murdoch y Cairo se realizaron por videollamada.

Kiko Llaneras es ingeniero y periodista de datos del diario español El País. El motivo de la entrevista responde a su experiencia en el periodismo de datos, especialmente en el desarrollo de modelos de predicción para elecciones políticas y eventos deportivos del mismo tipo que los que realiza FiveThirtyEight.

Por otra parte, John Burn-Murdoch es periodista de datos en el medio británico Financial Times. La razón de su entrevista es que es un periodista de datos que ha trabajado en periódicos con una fuerte apuesta por este tipo de periodismo como The Guardian o el propio Financial Times. Además, actualmente desarrolla modelos a través de lenguajes de programación para el desarrollo de historias periodísticas que se pueden narrar a través de los datos.

Daniele Grasso es el creador y coordinador de la unidad de periodismo de datos del medio digital español El Confidencial, que destaca por ser un diario innovador. Como miembro del Consorcio Internacional de Periodistas de Investigación (ICIJ), ha colaborado en la cobertura de investigaciones como los 'papeles de Panamá' o la revelación de cuentas secretas del banco HSBC.

Por último, Alberto Cairo es profesor en la Universidad de Miami, especializado en el área de visualización de datos, y doctor en periodismo por la Universitat Oberta de Catalunya. Además, es autor de varios libros especializados en periodismo de datos y su visualización.

Las entrevistas estaban estructuradas en torno a los siguientes temas:

- Situación actual del periodismo de datos.
- Motivos por los que el periodismo de datos está experimentando un crecimiento.
- Habilidades necesarias para la actividad del periodista de datos.
- Viabilidad de proyectos que pretendan utilizar el periodismo de datos como

nicho.

- Aspectos en los que se diferencia FiveThirtyEight de otros medios de comunicación que utilizan los datos.

Toda la información extraída de las entrevistas se ha integrado en el apartado de resultados, utilizándola como fuente para apoyar, oponer o comparar algunos de los puntos expuestos. En algunos casos, cuando las preguntas han ido referidas a la opinión de los expertos, algunas de las respuestas se han incluido cuando ha sido necesario dado que los entrevistados son expertos cuya opinión es relevante.

Además, también se han utilizado algunas técnicas cuantitativas. Desde el 18 de junio al 6 de julio, se recogieron y anotaron todos los artículos y entradas que FiveThirtyEight destacó en su página web. El objetivo de este registro era conocer cuáles son los principales temas tratados por el medio, así como la frecuencia de los mismos. También se han recopilado datos sobre el tráfico de audiencia en su sitio web y el volumen de trabajadores en la redacción en las distintas etapas del medio. Con esto se pretende comprobar, junto con algunas de las técnicas cualitativas, si un medio de este tipo es viable.

Esta información fue recopilada a través de distintas entradas de blog publicadas tanto por Nate Silver como por FiveThirtyEight. Los datos se anotaron en una tabla de una hoja de cálculo para establecer una comparativa de la evolución de los mismos.

## **4. Resultados**

### **4.1. Organización**

En este apartado, se muestra la situación actual de la redacción del medio. El objetivo es mostrar cómo está organizada, cuál es la distribución del personal, qué áreas se trabajan y cómo ha ido evolucionando. Para ello, hay que tener en cuenta que FiveThirtyEight ha tenido tres etapas claramente distinguidas que han provocado diferentes cambios en su estructura.

Su primera etapa se da como blog independiente, entre 2008 y 2010. Las elecciones norteamericanas de 2008 y el exitoso modelo predictivo de Nate Silver generaron la formación de este blog que originalmente consistía en la publicación de modelos probabilísticos enfocados a los resultados de los comicios políticos. La innovación de Silver con sus modelos de agregación de encuestas para elecciones estadounidenses e internacionales le hicieron ganar un gran prestigio.

Esto lleva a la segunda etapa, como parte de The New York Times. Silver cedió la licencia para que el blog formase parte del diario neoyorquino por tres años, mientras que le mantenían al frente del mismo. Aquí, se desarrolló una plantilla de siete redactores, aunque no todos publicaban con la misma regularidad y solo dos trabajaban para FiveThirtyEight a tiempo completo. El renovado éxito de Silver en la predicción de las elecciones de 2012 dio, de nuevo, gran prestigio al blog. Cabe destacar que su paso por The New York Times generó el producto que ahora es competencia de FiveThirtyEight: The Upshot.

En 2013, tras expirar la licencia cedida al medio neoyorquino, ESPN, el gigante de la comunicación deportiva en Estados Unidos, se hizo con los derechos del blog. FiveThirtyEight pasó entonces a tener una plantilla de 20 trabajadores y comenzó a publicar sobre cuatro grandes bloques además de política: deportes, economía, ciencia y cultura. Desde abril de 2018, FiveThirtyEight es parte de ABC News, una división de The Walt Disney Company, a la que también pertenece ESPN.

En la actualidad, la plantilla se compone de 35 personas. 17 son hombres (49%) y 18 son mujeres (51%). Por rangos, encontramos cuatro divisiones. Un redactor jefe, Nate Silver; cuatro directores de área, de los cuales dos son hombres y dos mujeres; cuatro editores jefe, tres de ellos hombres y una mujer; y 26 trabajadores con distintas

funciones entre las que se encuentran redactores, productores, coordinadores, editores y un desarrollador web. En esta última categoría, 11 son hombres y 15 mujeres.

En cuanto a la división por áreas, de los 26 trabajadores del rango más bajo, cinco están dedicados a la visualización, dos componen el área de política, dos el de deporte, dos el de ciencia y uno el de salud. Además, dos se dedican al desarrollo y a la producción del podcast semanal del medio y uno al formato audiovisual. Hay un informático encargado del desarrollo de la web y otros diez trabajadores no tienen una categoría específica asignada.

En cuanto al trabajo en equipo, en base a los artículos y entradas de la web analizados en el periodo de observación, se aprecia que el 65% llevan la firma de un único redactor, el 23% estaban realizados por dos o más escritores y tan solo el 12% no llevaban firma o se trataba de contenido editorial que iba firmado como “FiveThirtyEight”.

## **4.2. Contenidos**

Los contenidos de FiveThirtyEight se distribuyen en tres grandes bloques: artículos, modelos probabilísticos y la sección de audio y vídeo.

### **Artículos**

Las entradas de FiveThirtyEight son variadas en cuanto a su contenido. Por un lado, están los artículos, con un formato periodístico más tradicional, pero siempre con el apoyo visual de gráficos, infografías o presentación de datos en tablas. Durante el período observado, se encontraron cinco tipos de entradas diferentes:

- Artículos en base a uno o varios datos que eran la fuente principal de la noticia (conformaban el 60% de las entradas).
- Predicciones políticas y deportivas, con artículos en los que se explicaba cómo habían desarrollado los modelos (20%).
- Encuestas y sondeos sobre elecciones políticas (8%).
- Recopilaciones de los datos más importantes de la jornada a modo de resumen de actualidad (7%).

- Cuestionarios y juegos interactivos que se basaban en el *feedback* o retroalimentación de los usuarios (5%).

En cuanto a la temática de los contenidos, cabe destacar que la mayoría pertenece a la categoría de ‘política’ o a la de ‘deporte’. En base al período de observación, el 42% de las entradas trataban sobre deporte, mientras que el 38% lo hacían sobre política. Además, un 13% de los contenidos se encontraban en la categoría ‘cultura’ y un 7% lo hacía en ‘ciencia y salud’. Ninguno de los contenidos se encontraba en ‘economía’. Cabe destacar que, durante el mencionado período, se dieron eventos importantes que pudieron condicionar estos resultados tanto a nivel deportivo, como el Mundial de fútbol o las eliminatorias por los campeonatos de la NBA y la NHL; como a nivel político, con la celebración de elecciones primarias en algunos de los estados norteamericanos.

### **Modelos probabilísticos**

En un apartado distinto, FiveThirtyEight tiene destacados los modelos de predicción. Se trata únicamente de las predicciones, no de los artículos mencionados anteriormente en los que explica cómo se han llevado a cabo. Los modelos solamente abarcan las áreas de política y deporte y durante el período de observación, hubo cinco activos que se encontraban destacados:

- Predicciones sobre la Copa Mundial de Fútbol.
- Modelo de agregación de encuestas sobre la aprobación al presidente de Estados Unidos, Donald Trump.
- Predicciones sobre la liga estadounidense de béisbol profesional (MLB).
- Modelo de agregación de encuestas sobre las próximas elecciones al Congreso estadounidense.
- Predicciones sobre distintas ligas de fútbol profesional en Europa y América cuya competición estaba activa.

Estos modelos son los que han diferenciado a FiveThirtyEight desde su creación. Originalmente eran sobre política, pero tras el acuerdo con ESPN, se empezaron a desarrollar también sobre deporte. Para el apartado deportivo, se tomó como referencia el Elo, el sistema de puntuación utilizado para las clasificaciones en ajedrez. La base de

esta puntuación consiste en que cuanto mayor sea el rival al que se enfrenta un deportista o un equipo, más puntos puede conseguir por ganar. Y al contrario, cuanto peor sea el rival contra el que pierda, más puntos perderá por dicha derrota. El Elo se utiliza en FiveThirtyEight para la mayoría de deportes, que suelen ser de puntuación alta y no puede haber empates. La excepción es el fútbol, deporte para el que Nate Silver desarrolló un sistema de puntuación similar al Elo al que llamó Soccer Power Index (SPI). La diferencia que tiene respecto al Elo es que distingue entre ataque y defensa, y tiene en cuenta el bajo número de tantos que se pueden conseguir en esta disciplina.

TEAM ↕	DIVISION ↕	TEAM RATING ↕	1-WEEK CHANGE ↕	AVG. SIMULATED SEASON		POSTSEASON CHANCES		
				RECORD ↕	RUN DIFF. ↕	MAKE PLAYOFFS ↕	WIN DIVISION ↕	WIN WORLD SERIES ↕
<b>Red Sox</b> 94-44	AL East	1589	-7	109-53	+237	>99%	98%	25%
<b>Astros</b> 83-53	AL West	1589	-1	99-63	+251	>99%	85%	16%
<b>Cubs</b> 81-55	NL Central	1555	+7	96-66	+140	>99%	87%	12%
<b>Dodgers</b> 75-62	NL West	1562	+6	89-73	+145	72%	57%	11%
<b>Indians</b> 77-59	AL Central	1558	+1	93-69	+160	>99%	>99%	8%
<b>Yankees</b> 86-51	AL East	1565	-10	100-62	+164	>99%	2%	7%
<b>Braves</b> 76-60	NL East	1520	+2	89-73	+102	82%	80%	4%
<b>Athletics</b> 82-56	AL West	1553	-1	96-66	+100	94%	15%	4%
<b>Cardinals</b> 76-61	NL Central	1541	+3	89-73	+80	71%	5%	4%
<b>Diamondbacks</b> 74-63	NL West	1531	-1	87-75	+89	41%	23%	3%

Figura 1: Captura de pantalla de las predicciones para la Liga Mayor de Béisbol (MLB). Al lado de cada título aparecen dos pequeñas flechas triangulares que indican que esa categoría es ordenable. De esta manera, el usuario podría ordenar esta tabla en nueve categorías diferentes. Fuente: FiveThirtyEight.com

Los modelos probabilísticos que ofrece FiveThirtyEight en la categoría de deportes acostumbran a tener cierta interactividad. De esta manera, un usuario puede, por ejemplo, ordenar a los equipos según diferentes categorías como su Elo (o SPI), la probabilidad de que alcancen la siguiente ronda o la de que ganen el campeonato (figura 1). Además, en el caso de la Copa Mundial de fútbol, a la interactividad de los datos en tabla, FiveThirtyEight también permitía la posibilidad de completar el cuadro de eliminatorias. Para hacerlo, el usuario podía elegir al equipo que consideraba que avanzaría a la siguiente fase en base a la probabilidad de victoria que FiveThirtyEight le otorgaba, según su modelo, mostrada en porcentaje (figura 2). También durante el Mundial de fútbol, FiveThirtyEight introdujo por primera vez la predicción en directo,



que iba actualizando los porcentajes de victoria, empate y derrota minuto a minuto mientras el partido se estaba disputando (figura 3).

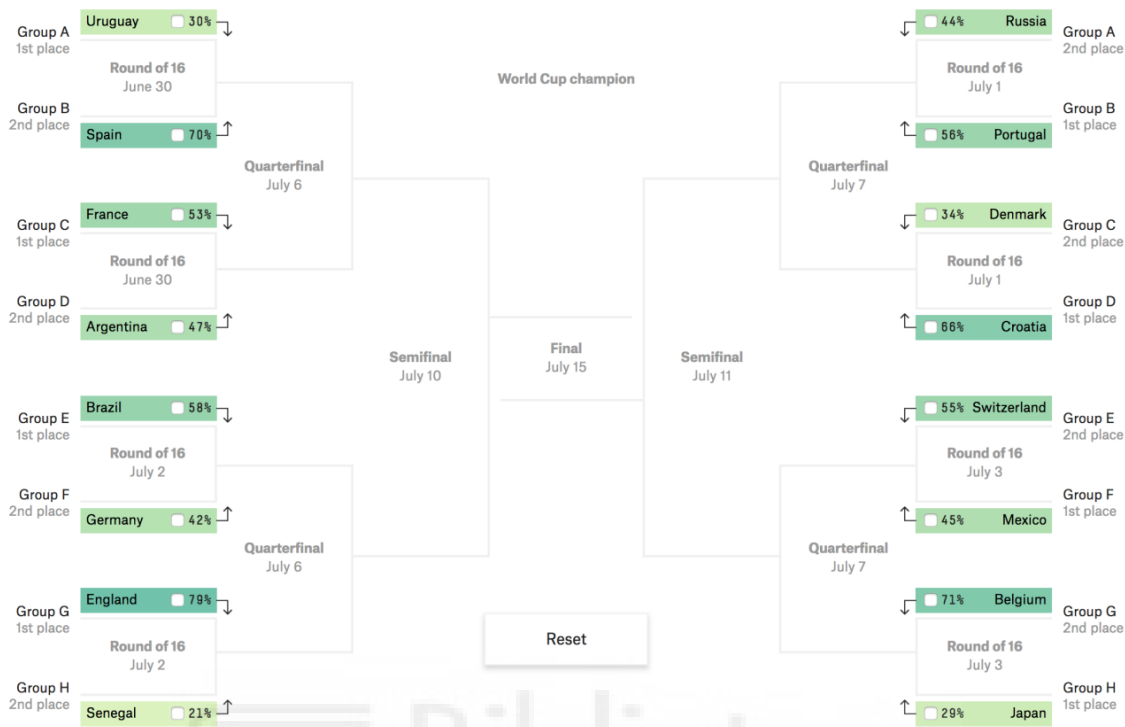


Figura 2: Captura de pantalla de las predicciones interactivas para las eliminatorias de la Copa Mundial de fútbol. En este modelo, aparece la probabilidad de victoria de cada equipo en cada emparejamiento. El usuario puede seleccionar quién cree que va a avanzar de ronda marcando al equipo ganador en los pequeños recuadros blancos. Fuente: FiveThirtyEight.com

Wednesday, June 27

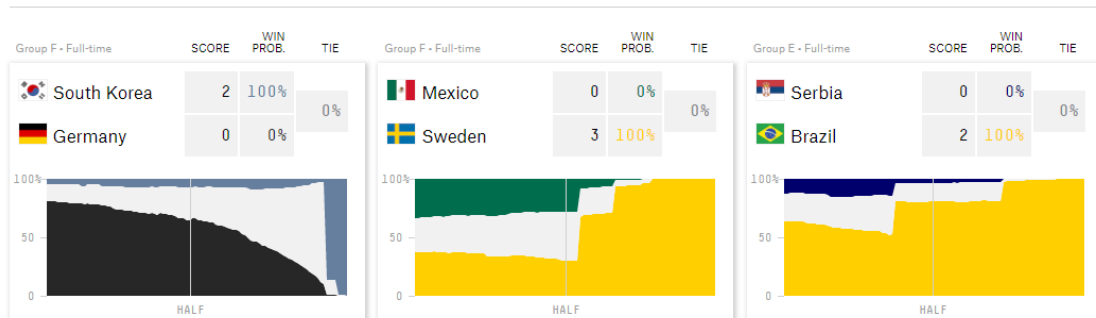


Figura 3: Captura de pantalla de las predicciones en directo (con el partido finalizado) de tres enfrentamientos del Mundial de fútbol. Cada uno de los colores representa la probabilidad de victoria de un equipo, mientras que el espacio en blanco es el empate. Los cambios bruscos se daban después de que un conjunto anotase un gol. Fuente: FiveThirtyEight.com

Los trabajos de predicción en la sección política consisten, principalmente, en el modelo de agregación de encuestas. Este pondera diferentes aspectos de diferentes sondeos sobre elecciones o actualidad política, y los muestra en un gráfico de líneas. La interacción, aunque mínima, permite ver los porcentajes en días anteriores o filtrar el tipo de encuestas que el usuario quiere que se muestren (figura 4).

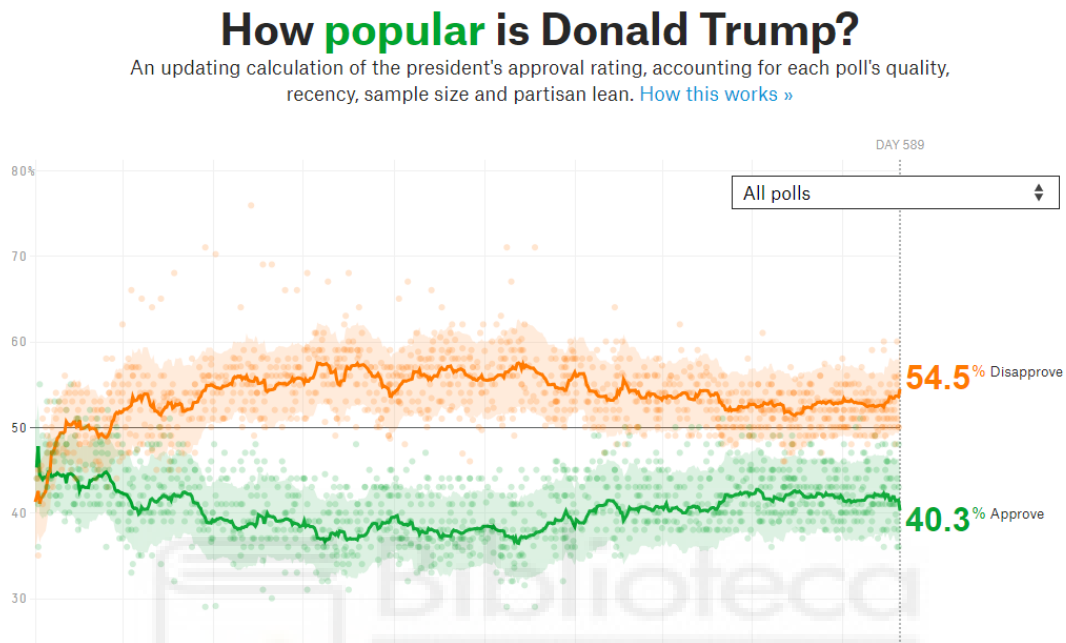


Figura 4: Captura de pantalla del modelo de agregación de encuestas sobre la aprobación al presidente de Estados Unidos, Donald Trump. Al mover el cursor sobre las líneas, muestra sobreimpreso el porcentaje en cada día. Además, en la caja blanca de la parte derecha, se pueden filtrar los resultados según el tipo de encuestas. En la descripción aparece un enlace que lleva al artículo donde está explicado el funcionamiento del modelo. Fuente: FiveThirtyEight.com

## Podcast y vídeo

Además de los artículos en texto, FiveThirtyEight también cuenta con secciones de audio y vídeo. En el caso del audio, el medio desarrolla un podcast semanal sobre política. La sección tiene más de 200 registros y, en total, el medio ha grabado más de 500 podcasts. Por otra parte, también cuenta con un apartado audiovisual, aunque ocupa el último espacio de la portada y no se ha apreciado regularidad en la actualización de este tipo de contenido durante la observación del medio. Ambos se desarrollan desde 2014, cuando se lanzó la plataforma en ESPN, y tienen una importancia menor respecto al resto de contenidos.

### 4.3. Innovación y rasgos distintivos

FiveThirtyEight es un medio único y pionero en su campo, por lo que aporta constantes innovaciones al periodismo de datos y al uso de los mismos. Desde su lanzamiento hace una década, los modelos probabilísticos han supuesto una novedad que le han convertido en referencia. El modelo de Nate Silver para las elecciones estadounidenses de 2008 supuso establecer una base en cuanto a la agregación de datos para las encuestas políticas. Para Kiko Llaneras, que en España utiliza modelos del mismo tipo, Silver se convirtió en el referente entre 2008 y 2012 por varios factores:

“El uso de los datos actual en periodismo es un fenómeno ligado a internet. Se empieza a poder hacer porque hay muchas encuestas, y luego hay gente capaz de desarrollar esto. Antes de internet, no tenía sentido crear un modelo de predicción solo para contárselo a tus conocidos. Internet y los blogs permiten mostrar contenidos a mucha gente. Silver llega desde un blog de frikis del béisbol. La gente empieza a ver este nuevo contenido y le gusta”.

Además, para Llaneras, la expectación generada en esas elecciones influyó para el especial seguimiento a un blog emergente como el de Silver.

“Silver también se ha convertido en la referencia porque estaba en Estados Unidos y las elecciones de ese país tienen una especial atención. Además, las de ese año tuvieron un interés añadido. Eso provocó que su modelo se fuera replicando”.

A esto hay que sumar que el periodismo en Estados Unidos tiene una capacidad económica superior al que se desarrolla en Europa. Como afirma Daniele Grasso: “Cuando algo se compara a Estados Unidos, hay que tener presente que se habla de otra escala de dinero”.

Los modelos predictivos de encuestas políticas han tenido un gran éxito y son la base de la innovación que promueve FiveThirtyEight. Para Llaneras, este éxito se debe a tres factores.

“En primer lugar, las preguntas son relevantes. Saber qué va a pasar o quién va ganando es importante. Tanto desde el punto de vista de la información como del punto de vista político, ya que los electores pueden modificar su voto en

base a esto. (...) En segundo lugar, la gente llega a los medios de comunicación para entretenerse. Se consumen, habitualmente, en el tiempo libre. En este sentido, estos modelos tienen una parte de retransmisión que es muy interesante, y es un producto que Nate Silver cuida especialmente bien. (...) Y en tercer lugar, un buen tratamiento de los números da confianza al lector. Los datos generan una información veraz para la audiencia.”

En la parcela del periodismo deportivo de datos, FiveThirtyEight también ha producido bastantes modelos innovadores. La aplicación del Elo a otras disciplinas que no son el ajedrez ha dado lugar a distintos tipos de modelos de predicción o probabilísticos. Algunas de las innovaciones más destacadas que ha aportado FiveThirtyEight en este apartado son:

- El uso del sistema de puntuación Elo para deportes distintos al ajedrez.
- La creación del sistema de puntuación Soccer Power Index (SPI), similar al Elo pero específico para el fútbol.
- El uso del Elo para comparar a jugadores de deportes de equipo. El más usado es el CarmElo (nombrado así por el baloncestista Carmelo Anthony), que predice el rendimiento que tendrá el atleta en las siguientes siete temporadas en base a su rendimiento previo y a la comparación con otros deportistas de las mismas características.

Para Kiko Llaneras, los principales motivos que llevan al uso del Elo o de modelos creados en base a este sistema son la sencillez a la hora de calcularlo y sus propiedades matemáticas, las cuales son “útiles para la predicción”. Lo que hace al Elo especialmente interesante para el ámbito deportivo, según Llaneras, es el hecho de que permite comparar a dos rivales aunque estos lleven bastante tiempo sin enfrentarse, ya que el Elo tiene en cuenta los enfrentamientos con equipos de diferente nivel.

Para Daniele Grasso, la clave de FiveThirtyEight también está relacionada con la aplicación de los modelos de datos: “(Lo que diferencia a FiveThirtyEight) es que vuelcan todo, absolutamente todo, desde un punto de vista de ciencia de datos. Su apuesta es aplicar los métodos científicos de datos para crear historias periodísticas”.

Por el contrario, el periodista Burn-Murdoch cree que, aunque estos modelos fueron

innovadores, ahora no lo son tanto y no se distinguen por ellos.

“No creo que hagan nada que sea muy distinto a lo que realizan otras redacciones. (...) Aquello en lo que más destaca que es el análisis político y los modelos de predicción también se ve ahora en otros medios como The New York Times o The Economist. Los modelos de FiveThirtyEight son realmente complejos y es un trabajo muy bueno, pero incluso en el Financial Times estamos empezando a utilizar técnicas similares”.

En este punto discrepa Ritchie King, coordinador del área de visualización de FiveThirtyEight, quien considera que aunque otros medios trabajan con este tipo de modelos, 538 sigue innovando en ellos, lo que les hace ir por delante:

“Dentro de los modelos de predicción, estamos comenzando a usar predicciones en directo y a utilizar la jugabilidad. Ahora mismo, estamos intentando hacer llegar a la gente datos en tiempo real. (...) Intentamos ir un paso por delante, no sólo dar un pronóstico de los resultados, eso es lo que hace que sigamos innovando”.

### **La figura de Nate Silver**

Pese a que actualmente la plantilla se compone de más de una treintena de personas y que los contenidos no son generados solamente por Nate Silver, como sucedía hasta 2012, su presencia en el medio como redactor jefe sigue siendo una de las claves del éxito y de la continua innovación de FiveThirtyEight.

Margareth Sullivan, quien en 2013 era editora jefe en The New York Times, publicó lo siguiente acerca de la figura de Silver tras su marcha a ESPN.

“No creo que Nate Silver encajase en la cultura del Times y creo que él se dio cuenta. Ha sido, en una palabra, disruptivo. Nate Silver ha cambiado por completo el modo de cubrir la actualidad política” (Sullivan, 2013).

En el mismo artículo, Sullivan también hablaba acerca del choque que provocó esta forma de tratar la política con periodistas tradicionales del diario neoyorquino.

“Toda su forma de ver la política basada en la probabilidad iba en contra del periodismo político en el que The Times se especializa. (...) Su enfoque fue trabajar en contra de la narrativa de la política, y eso hizo que leer su contenido

fuera interesante. Para mí, ambos enfoques tienen valor y pueden convivir. (...) A varios periodistas tradicionales les disgustó su trabajo”.

Además, Sullivan también destacó la capacidad que tuvo Silver de crear su propio público.

“Silver tiene poder de estrella. Su capacidad de generar tráfico, especialmente entre lectores jóvenes que no suelen leer el periódico, no ha tenido comparación, y seguramente seguirá siendo así”.

En este punto, Sullivan coincide con la opinión de Kiko Llaneras, quien también destaca la figura de Silver como principal rasgo distintivo de FiveThirtyEight frente a otros productos similares:

“Lo que diferencia a FiveThirtyEight es que tiene a Nate Silver. Es una persona muy intuitiva con los modelos y los ha diseñado con una metodología que personalmente me gusta. Su tratamiento de los datos es muy bueno, tiene un buen tono y una plantilla de gente muy talentosa”.

Para Alberto Cairo, la figura de Nate Silver como jefe es fundamental porque supone que es un experto en periodismo de datos el que confecciona la plantilla, lo que le da ventaja sobre la competencia:

“Desde el principio la gente que trabajaba en FiveThirtyEight era especialista en datos y estadística. Los medios tradicionales para hacer este tipo de contenido tienen que contratar a gente nueva, pero la gente que contrata a otra en FiveThirtyEight es gente que ya trabaja con este periodismo, como Nate Silver. En FiveThirtyEight, Nate Silver es el jefe, y los subdirectores seguramente también son frikis de los números. Incluso los pocos medios tradicionales que han ascendido a especialistas en datos a puestos de poder, como The New York Times y el Washington Post, están en desventaja”.

#### **4.4. Audiencia**

Pese a que en 2016 FiveThirtyEight elaboró el artículo más leído del año, uno de los factores clave para entender su éxito, es que su audiencia está especializada y segmentada, algo que para Alberto Cairo, es un factor diferencial:

“La (audiencia) de FiveThirtyEight es diferente. Los lectores de este medio son lo que en Estados Unidos se conoce como *wonks*, los frikis. Al que le gusta la sección de deportes de FiveThirtyEight seguramente le gustan tanto los números del deporte como el deporte en sí”.

El modelo de medio actual de FiveThirtyEight es el usado desde su lanzamiento como parte de ESPN en 2014. En aquel momento, la plantilla pasó de tener dos redactores a tiempo completo a tener veinte, y los contenidos pasaron de estar centrados en política a cubrir las cinco grandes áreas que el medio sigue tratando actualmente: política, deportes, ciencia y salud, cultura y economía (Silver, 2014).

Esta apuesta de multiplicar por diez el tamaño de la redacción, se vio reflejada en los resultados de la audiencia. En los siguientes dos años, entre 2014 y 2016, el número de usuarios únicos mensuales pasó de 2,8 a 10,7 millones, según Omniture. Si bien es cierto que 2016 fue un año electoral en Estados Unidos y los modelos de predicción de Silver son especialmente seguidos, en 2015, la audiencia al cierre del ejercicio superaba los 8 millones de usuarios únicos.

Otro de los rasgos que distinguen a FiveThirtyEight es su capacidad para desarrollar productos de nicho que no tendrían cabida en otro medio, algo que destaca Llaneras.

“FiveThirtyEight hace cosas que no interesan a la suficiente gente. Es normal porque su audiencia recoge a todas las personas interesadas en este tipo de información que no se encuentra en otros medios de comunicación, pero esto también se debe a que cuando se plantea llevar algo que ellos hacen a nuestra audiencia (en El País), tenemos que preguntarnos si tiene sentido para nuestro público, y no siempre es así”.

Este es un punto en el que coincide con Daniele Grasso, quien también considera que FiveThirtyEight está excesivamente especializado en un tipo de audiencia a la que no es fácil llegar.

“A mí FiveThirtyEight, salvo algunas cosas concretas que son absolutamente espectaculares, me parece que no es un periódico por el que yo pagaría por leer su contenido. Están muy especializados y han encontrado ese filón”.

Por su parte, Burn-Murdoch pone en duda por su experiencia que se pueda llevar a cabo

un proyecto como el de FiveThirtyEight en un medio como el Financial Times.

“Cuanto más lo pensamos y analizamos (crear proyectos basados en los datos en el Financial Times), nos damos cuenta de que la gran mayoría de los lectores están interesados en un tema concreto, ya sea economía, política o medio ambiente; y ya sea en texto, *podcast* o gráfico lo que quiere el lector es informarse sobre esos temas. Normalmente un lector no nos pide que le enseñemos cualquier gráfico o mapa sino aquellos que están relacionados con el tema del que quieren saber más”.

Ritchie King afirma que algo que diferencia a FiveThirtyEight es que no crean los contenidos en base al volumen de audiencia:

“No estamos tan pendientes de las analíticas como para estar pensando continuamente si un contenido ha funcionado mejor que otro. Nuestros principales objetivos son la calidad y la precisión”.

Todos estos factores en conjunto, hacen de FiveThirtyEight un medio con una audiencia única, especializada y formada. En opinión de Ritchie King, esto permitiría al medio sobrevivir de manera independiente a un gran medio de comunicación:

“Creo que FiveThirtyEight sería viable por sí mismo. Podría funcionar bajo suscripción, al menos una parte de lo que hacemos como los modelos de predicción. Nuestro público suele ser gente con cierto nivel de formación y entendería que parte de lo que hacemos tuviera un muro de pago”.



## 5. Conclusiones

Tras el análisis de los resultados, en esta parte se expondrán las conclusiones. El objetivo principal era conocer cuáles son los contenidos y estrategias que utiliza FiveThirtyEight para desarrollar su actividad como medio de comunicación basado en el periodismo de datos.

Por un lado, hay que destacar el uso de los modelos probabilísticos tanto en política como en deporte. Son la seña de identidad desde que se fundó 538 hace una década y todavía hoy se sigue innovando en ese aspecto, desarrollándolos para eventos como competiciones deportivas y encuestas políticas, a los que se añaden funciones como la de jugabilidad.

Sin embargo, este tipo de contenido es cada vez más habitual en otros medios de comunicación, especialmente en tradicionales como The New York Times. En España ya se desarrollan en medios como El País, e incluso periódicos cuyo foco no es la política ni el deporte como es el británico The Financial Times, especializado en economía, están comenzando a utilizar técnicas similares.

Pese a ello, FiveThirtyEight sigue siendo referencia en este campo. La figura de Nate Silver es, según Margareth Sullivan y Kiko Llaneras, clave para entender esto. Por un lado, es una figura que atrae a un gran volumen de audiencia por sí mismo. Por otro, su intuición a la hora de desarrollar los modelos genera una renovación constante de los mismos, siendo la más reciente la introducción de los pronósticos en directo para eventos deportivos.

Otro de los objetivos que se planteaban de cara a la realización de este trabajo era concluir si es viable un medio que utilice los datos como nicho o si, por el contrario, se trata de un producto de prestigio para grandes medios, ya que FiveThirtyEight siempre ha estado ligado a uno.

Aunque tiene una audiencia de diez millones de usuarios únicos mensuales, no hay cifras sobre el tiempo de estancia en la web ni sobre el número de entradas que cada usuario visita. Según Llaneras, “FiveThirtyEight a veces peca de friki”, una opinión que coincide con la de Daniele Grasso, que aunque reconoce ser consumidor de datos, no pagaría una suscripción a un medio como 538 por ser demasiado especializado. Para Burn-Murdoch, la audiencia no busca la información en un formato concreto, sino en

los temas que le interesan. Por su parte, Alberto Cairo entiende que FiveThirtyEight cuenta con la ventaja de tener una audiencia sesgada que le permite hacer un contenido especializado sin riesgo.

Aunque Ritchie King, de FiveThirtyEight, cree que podrían funcionar de manera independiente añadiendo muro de pago en algunas secciones o contenido patrocinado, se intuye que FiveThirtyEight es, efectivamente, un producto de prestigio para un gran medio de comunicación como ahora mismo es ABC News. Algunos de sus contenidos tienen gran importancia, como la predicción a las elecciones norteamericanas de 2016, que fue la información más leída de ese año a nivel mundial, pero la mayoría de ellos son demasiado especializados como para llegar al gran público.

Por otra parte, encontramos que FiveThirtyEight tiene ventajas culturales y económicas al estar establecido en Estados Unidos. En cuanto a las ventajas culturales, encontramos que las elecciones de su país tienen gran relevancia a nivel mundial. En el ámbito deportivo, otro de los más destacados, algunos de los deportes mayoritarios de EE. UU. como el béisbol y el baloncesto son más medibles mediante las estadísticas que, por ejemplo, el fútbol europeo. Además, hay que sumar que Estados Unidos es un país con más de 300 millones de habitantes que utiliza la lengua inglesa, hablada por, aproximadamente, mil millones de personas. En el apartado económico, el periodismo norteamericano cuenta con más recursos que el de otras zonas como la europea, según Grasso, por lo que un proyecto de estas características puede llegar a contar con un volumen de trabajadores que, por el momento, es impensable en Europa.

Los mayores retos a los que se enfrenta ahora FiveThirtyEight son varios. Por un lado, los modelos predictivos tienen el desafío de explicar a los lectores los resultados de sus predicciones. Una dificultad que quedó patente tras las elecciones de 2016, en las que Donald Trump se hizo con la presidencia pese a que en el modelo solo ganaba en un tercio de las simulaciones, algo que fue comprendido por muchos lectores como un error.

Por otra parte, podría perder importancia si otros medios siguen desarrollando los mencionados modelos. Sin embargo, esta posibilidad parece lejana, ya que su continua innovación con nuevos modelos como la proyección de rendimiento de los jugadores de la NBA (CarmElo) dificulta que puedan perder esta posición de referencia.

## 6. Notas

1. Las *sabermétricas*, del término inglés *sabermetrics*, acuñado por el periodista y escritor norteamericano Bill James y que deriva de las siglas de la Society for American Baseball Research (SABR); son los modelos utilizados en béisbol para estudiar y predecir las diferentes situaciones de juego que se producen en los partidos mediante el uso de las estadísticas. A los analistas que utilizan estas técnicas se les conoce como *sabermétricos* o, en inglés, *sabermetricians*.
2. 538 fueron los votos electorales que se disputaron en las elecciones presidenciales estadounidenses de 2008. El nombre de la web FiveThirtyEight.com procede de ahí. En las últimas elecciones, las de 2016, el número de votos electorales fue de 531.



## 7. Bibliografía

- **Carlson, M. & Usher, K.** (2015). “News Startups as Agents of Innovation”. *Digital Journalism*, 4(5), 563-581. <https://www.tandfonline.com/doi/full/10.1080/21670811.2015.1076344>
- **Chua, R.** (2010). “Structured journalism”. *(Re)structuring Journalism*. 12 de agosto. <https://structureofnews.wordpress.com/2010/08/10/whats-it-all-about>
- **CIS** (2015). “Barómetro de abril 2015”. *Centro de Investigaciones Sociológicas*. n. 3080, abril de 2015. [http://www.cis.es/cis/export/sites/default/-Archivos/Marginales/3080\\_3099/3080/es3080mar.pdf](http://www.cis.es/cis/export/sites/default/-Archivos/Marginales/3080_3099/3080/es3080mar.pdf)
- **Crucianelli, S.** (2012). “Introducción al periodismo de datos”. *Ijnet: Red de periodistas internacionales*. <http://ijnet.org/es/blog/claves-para-entender-el-periodismo-de-datos>
- **De-Lara-González, A., Árias-Robles, F., Carvajal-Prieto, M., & García Avilés, J. A.** (2018). “Ranking de innovación periodística 2014 en España. Selección y análisis de 25 iniciativas”. *El profesional de la información*, v. 24, n. 3, pp. 235-245. <http://dspace.umh.es/bitstream/11000/4573/1/7-Articulo%20ranking%202015%20EPI.pdf>
- **Felder, A.** (2009). “Case study: The FiveThirtyEight.com Predictive Model of the 2008 Presidential Election”. *STATS Magazine*, n. 50, pp. 3-9. ISSN 1053-8607.  
Recuperado de: [https://www.causeweb.org/cause/archive/stats/STATS\\_50.pdf](https://www.causeweb.org/cause/archive/stats/STATS_50.pdf)
- **Fink, K. & Anderson, C.** (2015). Data Journalism in the United States: Beyond the “usual suspects”. *Journalism Studies*, 16(4), 467-481.
- **Freixa, P., Pérez-Montoro, M. & Codina, L.** (2017). “Interacción y visualización de datos en el periodismo estructurado”. *El profesional de la información*, v. 26, n. 6, pp. 1076-1090. <https://doi.org/10.31145/epi.2017.nov.07>
- **Gray, J., Chambers, L., & Bounegru, L.** (2012). *The data journalism handbook: How journalists can use data to improve the news*. O'Reilly Media, Inc. Recuperado de: <https://books.google.es/books?hl=es&lr=&id=W-sHp0ECqPUC&oi=fnd&pg=PR5&dq=gray+chambers+bounegru+the+data+jou>

[rnalism+handbook&ots=\\_XKxIhTdqw&sig=UskdWsARq69B7PgouNBqXhVp  
eHQ#v=onepage&q=gray%20chambers%20bounegru%20the%20data  
%20journalism%20handbook&f=false](http://www.journalismhandbook.com/2013/09/06/gray-chambers-bounegru-the-data-journalism-handbook&ots=_XKxIhTdqw&sig=UskdWsARq69B7PgouNBqXhVpeHQ#v=onepage&q=gray%20chambers%20bounegru%20the%20data%20journalism%20handbook&f=false)

- **Holovaty, A.** (2006). “A fundamental way newspaper sites need to change”. *Adrian Holovaty blog*, 6 de septiembre. <http://www.holovaty.com/writing/fundamental-change/>
- **Holovaty, A.** (2008). “In memory of Chicagocrime.com”. *Adrian Holovaty blog*. 31 de enero. <http://www.holovaty.com/writing/chicagocrime.org-tribute/>
- **Holovaty, A.** (2009). “The definitive, two-part answer to ‘is data journalism’”. *Adrian Holovaty blog*. 21 de mayo. <http://www.holovaty.com/writing/data-is-journalism/>
- **Howard, A.** (2014). “The art and science of data-driven journalism”. *Tow Center for Digital Journalism Blog*. 30 de mayo. <http://towcenter.org/blog/the-art-and-science-of-data-driven-journalism/>
- **Lewis, S. & Westlund, O.** (2015). “Big data and journalism: Epistemology, expertise, economics and ethics”. *Digital journalism*, v. 3, n. 3, pp. 447-466. Recuperado de <https://apps.cla.umn.edu/directory/items/publication/324155.pdf>
- **López-García, X., Toural-Bran, C. & Rodríguez-Vázquez, A.** (2016). “Software, estadística y gestión de bases de datos en el perfil del periodista de datos”. *El profesional de la información*, v. 25, n. 2, pp. 286-294. <http://dx.doi.org/10.3145/epi.2016.mar.16>
- **McDuling, J.** (2014). “The Upshot is the New York Times’ replacement for Nate Silver’s FiveThirtyEight”. *Quartz*. 10 de marzo. <https://qz.com/185922/the-upshot-is-the-new-york-times-replacement-for-nate-silvers-fivethirtyeight/>
- **Meyer, P.** (1991). “The new precision journalism”. *Indiana Univ Pr*. Recuperado de: <https://www.unc.edu/~pmeyer/book/>
- **Niles, R.** (2006). “The programmer as journalist. A Q&A with Adrian Holovaty”. *Online Journalism Review*. 5 de junio. <http://www.ojr.org/the-programmer-as-journalist-a-qa-with-adrian-holovaty/>

- **Peiró, K. & Guallar, J.** (2013). “Introducció al periodisme de dades. Característiques, panoràmica i exemples”. *Ítem*, n. 57, pp. 23-37. <http://eprints.rclis.org/28670/>
- **Perry, F. & Paz, M.** (2014). “Manual de periodismo de datos iberoamericano”. *Fundación Poderomedia*. <http://manual.periodismodedatos.org/libro.php>
- **Royal, C.** (2012). “The Journalist as Programmer: A Case Study of The New York Times Interactive News Technology Department”. *ISOJ Journal*, v. 2, n. 1, pp. 5-24. [https://isoj.org/wp-content/uploads/2016/10/ISOJ\\_Journal\\_V2\\_N1\\_2012\\_Spring.pdf](https://isoj.org/wp-content/uploads/2016/10/ISOJ_Journal_V2_N1_2012_Spring.pdf)
- **Royal, C. & Blasingame, D.** (2015). “Data journalism: An explication”. #ISOJ, *the official research journal of the International Symposium on Online Journalism*, v. 5, n. 1, pp. 24-46. Recuperado de: <https://isojournal.wordpress.com/2015/04/15/data-journalism-an-explication/>
- **Silver, N.** (2008). “FAQ and Statement of Methodology. Last revised August 7th, 2008”. *FiveThirtyEight.com*. <https://fivethirtyeight.com/features/frequently-asked-questions-last-revised/>
- **Silver, N.** (2014). “What the Fox Knows”. *FiveThirtyEight.com*. 17 de marzo. <https://fivethirtyeight.com/features/what-the-fox-knows/>
- **Sullivan, M.** (2013). “Nate Silver Went Against the Grain for Some at The Times”. *The New York Times*. 22 de julio. <https://publiceditor.blogs.nytimes.com/2013/07/22/nate-silver-went-against-the-grain-for-some-at-the-times/>
- **Sullivan, M.** (2017). “Polls show Americans distrust the media. But talk to them, and it’s a very different story”. *The Washington Post*. 28 de diciembre. <https://www.washingtonpost.com/lifestyle/magazine/polls-show-americans-distrust-the-media-but-talk-to-them-and-its-a-very-different-story/2017/12/27/>
- **Wilson, A. & Katz, J.** (2013). “How Y’All, Youse and You Guys Talk”. *The New York Times*. <https://www.nytimes.com/interactive/2014/upshot/dialect-quiz-map.html>
- **Wojcik, L.** (2015). “Structured journalism, the next revolution in storytelling?”.

*Medium*, 4 October. <https://medium.com/interactive-journalism/structured-journalism-the-next-revolution-in-storytelling-91557c48c121>



## 8. Anexos

Las entrevistas a Ritchie King y John Burn-Murdoch fueron realizadas en inglés y las transcripciones fueron traducidas al castellano con la ayuda de Cristina Berenguer Pérez, graduada en Traducción e Interpretación por la Universidad de Alicante.

### 8.1. Entrevista a Ritchie King

**¿Qué factores hacen que el periodismo de datos sea una especialidad tan popular actualmente?**

El periodismo de datos se ha extendido mucho por internet, por las iniciativas de transparencia de los gobiernos y porque es más fácil que nunca utilizarlo. Hay mucha gente publicando datos en internet. Actualmente es muy fácil descargar una información y analizarla. Eso por una parte. Por otra, las herramientas para analizar y visualizar datos son mucho mejores, se pueden hacer más cosas con los datos que las que se podían hacer en el pasado. Los periodistas están aprovechando estas oportunidades.

**¿Qué habilidades necesitan los periodistas de datos? Supongo que en FiveThirtyEight la mayoría de ustedes son programadores.**

Lo más básico del periodismo de datos es, seguramente, algo tan simple como saber abrir una hoja de cálculo, tratar la información y escribir un texto sobre el resultado. Es un requisito mínimo, igual que tener gusto por las matemáticas. A partir de ahí se pueden hacer muchas cosas, desde *web scraping* (extraer datos de webs de forma automatizada); recolección de datos, que requiere algo de experiencia en programación; visualización, que es a lo que yo me dedico; modelos predictivos, que es lo que nos diferencia... Así que creo que el espectro es muy amplio y cada apartado del periodismo de datos necesita de una serie de herramientas diferentes.

**¿El hecho de que los programas que se utilizan sean más accesibles y en algunos casos gratuitos mejora el periodismo de datos?**

Sí, totalmente. Es más fácil de hacer y además es más fácil explicar el trabajo realizado a los lectores, lo que me parece muy importante.

**Dado que su campo es el de visualización, me gustaría saber por qué cree que se está abandonado el uso de la interactividad y hacia dónde se está dirigiendo este área.**



No diría que se está abandonando la interactividad. Creo que cuando se empezó a utilizar, la gente alucinó con las posibilidades que ofrecía y creyó que había que utilizarla siempre, pero luego empezó a darse cuenta de que era muy trabajoso para el lector y que no siempre era la mejor forma de contar una historia a través de los datos. Ahora, la interactividad se limita a los casos en los que es necesaria para contar una historia de la mejor manera. Es importante tener en cuenta que cualquier proyecto que se lleve a cabo hoy en día tiene que ser utilizable en los teléfonos móviles.

Además, creo que ha cambiado el concepto del periodismo de datos, no hace mucho la idea era enseñar todos los datos que se habían encontrado al lector, mientras que ahora la idea va más por el camino de mostrar las historias que nos parecen interesantes. El periodista filtra lo que considera importante. Todo esto afecta al uso de la interactividad. En nuestro caso, actualmente estamos utilizando la interactividad en cuestionarios y juegos, sobre todo, y de momento vamos a seguir apostando por ello.

**El contenido de FiveThirtyEight se basa exclusivamente en datos. ¿A qué fuentes acuden para encontrar temas?**

En política, nuestras principales fuentes son las agencias de encuestas. Además, están los datos electorales, que viene de fuentes oficiales como los estados estadounidenses. En cuanto a las encuestas, hay que tener algo en cuenta, cuando uno llama a una fuente para preguntarle sobre un tema, espera que la fuente sea experta en él. Para hacerlo, miras su trabajo y sus credenciales. Con las fuentes de datos eso es más complicado. Las gubernamentales son fiables, pero para las otras hay que ponderar su trabajo, en aspectos como su precisión en otras encuestas.

**Algunos periodistas de datos europeos, comentan que es más fácil desarrollar el trabajo con deporte en Estados Unidos porque sus disciplinas más populares como el béisbol o el baloncesto son más fáciles de cuantificar que el fútbol. Dado que ustedes también trabajan con fútbol, ¿está de acuerdo o cree que todos los deportes son medibles?**

La verdad es que no sé por qué se considera que el fútbol no es tan medible con números como otros deportes. Me parece que del fútbol hay datos interesantes como saber las zonas que ocupan los jugadores en el campo o comparar a dos jugadores de la misma posición para ver quién se mueve más por la zona ofensiva que por la defensiva.

Entiendo que tal vez la dificultad está en que hay muy pocas anotaciones, otros deportes tienen más puntos o goles. Realmente no soy un experto en deportes, me dedico más a la sección de política, pero creo que no ha habido una revolución a la hora de usar los datos en fútbol como sí la hubo en béisbol o en baloncesto. Al fútbol americano también le costó y aún no utiliza tantos datos como los otros, así que supongo que el siguiente deporte será el fútbol.

**FiveThirtyEight siempre ha funcionado con el apoyo de un medio de comunicación como ahora es ABC News. ¿Podría sobrevivir de manera independiente? ¿Es un proyecto viable o se trata de una marca de prestigio para una gran compañía?**

Tal vez es un poco vanidoso responder a esto, pero sí, creo que FiveThirtyEight sería viable por sí mismo. Podría funcionar bajo suscripción, al menos una parte de lo que hacemos como los modelos de predicción. Nuestro público suele ser gente con cierto nivel de formación y entendería que parte de lo que hacemos tuviera un muro de pago. También tenemos grandes proyectos que podrían ir patrocinados como los que hacemos en deportes. Es una pregunta difícil de responder, pero yo creo que sí que podríamos funcionar de manera independiente.

**La seña de identidad de FiveThirtyEight han sido los modelos de predicción, pero ahora otros medios también los utilizan. ¿Qué es lo que les diferencia ahora?**

Yo creo que todavía son los modelos de predicción. Quiero decir, The Upshot tiene un modelo de predicción para las elecciones pero nosotros tenemos una variedad mayor. Hacemos modelos para competiciones deportivas y, en general, creo que son más detallados que los de otros medios. Además, hacemos pronósticos en directo, algo que ha tenido un gran éxito ahora con el Mundial de fútbol, y también tenemos algo de jugabilidad. En la última temporada de NFL (fútbol americano) incluimos apartados como uno que consistía en que los usuarios tenían que intentar vencer a nuestro modelo. También podían adelantarse al resultado, por ejemplo, podían ver cómo el resultado de un equipo en la siguiente jornada, iba a afectar a las posibilidades del resto de la temporada. Intentamos ir un paso por delante, no sólo dar un pronóstico de los resultados, eso es lo que hace que sigamos innovando.

**¿Es The Upshot el mayor competidor de FiveThirtyEight?**

Probablemente (duda). Seguramente sean los que hacen un trabajo más parecido al

nuestro. En política seguro que sí, lo son. En deportes no, ahí no tenemos esa especie de rival.

**Pese a que en la página principal de FiveThirtyEight se destacan cinco áreas, al final la mayoría del contenido está basado en política y deporte. ¿Por qué pasa esto?**

Bueno, cuando FiveThirtyEight comenzó, era el proyecto de Nate Silver y tanto política como deporte son sus áreas. Cuando nos unimos a ESPN, quisimos ampliar los temas a cultura, economía y ciencia. Pero siempre supimos que iba a ser política y deporte en el escalón de arriba y los otros tres temas en el de abajo. Actualmente seguimos explorando formas de utilizar de manera correcta el periodismo de datos en estas tres áreas, seguramente en la que más éxito hemos tenido ha sido en ciencia. Se podría decir que ahora mismo estamos haciendo política, deporte y ciencia, son nuestras áreas principales.

**¿En qué está tratando de innovar FiveThirtyEight dentro del periodismo de datos?**

Seguramente en lo que comenté anteriormente, dentro de los modelos de predicción, estamos comenzando a usar predicciones en directo y a utilizar la jugabilidad. Ahora mismo, estamos intentando hacer llegar a la gente datos en tiempo real con los que pueden interactuar, seguramente es nuestro mayor reto.

**Cuando se plantea una nueva propuesta para realizar en FiveThirtyEight, ¿tienen en cuenta que vaya a ser algo con mucha audiencia o si piensan que es una buena idea la sacan adelante igualmente?**

Creo que algo que nos diferencia de otros medios es que no estamos tan pendientes de las analíticas como para estar pensando continuamente si un contenido ha funcionado mejor que otro. Nuestros principales objetivos son la calidad y la precisión. Para ello tenemos unos editores muy comprometidos, revisamos mucho la información que obtenemos e intentamos ser muy cuidadosos. Nuestro principal objetivo es hacer las cosas bien. Creo que eso nos da posibilidades para experimentar y ser creativos, lo cual es genial, porque estamos en un área que tiene muchas posibilidades para innovar y probar cosas nuevas.

## **8.2. Entrevista a Kiko Llaneras**

### **¿Por qué tienen tanta relevancia los datos en el periodismo actual?**

Es verdad que la página de FiveThirtyEight en la que Nate Silver siguió las elecciones de 2016 fue, según ChartBit, la noticia más leída del mundo en ese año. Esas elecciones fueron las más seguidas en todo el mundo y es una pieza recurrente. ¿Por qué hay ese interés por los datos en prensa? Yo creo que se dan dos cosas, una cuestión de oferta y una de demanda.

Por un lado, hay una mayor oferta de datos por la transformación que ha sufrido la sociedad que nos ha llevado a tener una mayor información cuantitativa, de *datasets*, que antes no existía. Tenemos más datos a nuestro alcance, los periódicos tienen más datos a su alcance y por lo tanto se pueden hacer más temas con esta orientación.

Del lado de la demanda, yo creo que hay una cuestión social y en cierto modo generacional. Hay un público, o una parte de él, que está más familiarizado con el uso de números. Se sabe manejar con estadísticas porque lo hace en su trabajo, tiene una formación de ciencias o simplemente porque jugaba al FIFA o al Pro Evolution Soccer y sabe leer y manejar los datos desde niño. Esta generación es ahora lectora de periódicos y demanda esta información cuantitativa. Así que yo creo que se juntan las dos cosas.

Luego hay un tercer elemento y es que los datos no son cosa del periodismo, el periodismo llega algo tarde. La cuantificación y la llegada de los números a un montón de industrias es anterior. En cierto modo en el periodismo hemos llegado tarde, pero si ahora vas a cualquier gran empresa la mayoría de la gente de esa empresa está manejando números.

### **¿Tiene influencia para la llegada de los números al periodismo que, según diversas encuestas, los ciudadanos hayan perdido la confianza en los medios?**

Es parte de la ventaja de los datos, aunque puede ser una maldición también. Los números tienen una aureola de confianza. Yo creo que esto está bien y está justificado. No es que sea imposible mentir con un dato o mentir con estadísticas, pero es más difícil ya que es más fácil comprobar que alguien te está diciendo algo que no es. Es algo verificable. Si yo digo que un partido político va a tener en torno al 25 o 30% de

votos y luego tiene el 50, me puedes pasar cuentas. O me puedes decir: “tú me estás diciendo esto pero en este otro sitio dicen esta otra cosa”. Es más fácil de verificar.

Si en lugar de eso yo digo cosas como que un partido político va muy fuerte y podría dar la sorpresa, eso es inverificable. En el fondo, no es una predicción. Si al final resulta que ganan yo puedo decir aquello de “ya lo dije”, pero si no pasa también se justifica, “yo no dije que fuera a ganar”. Son aseveraciones no tan falsables y venimos un poco de ese mundo.

### **¿Por qué tienen tanto éxito los modelos predictivos sobre elecciones políticas?**

Hay varias razones. La primera es que son preguntas importantes. El saber qué va a pasar, saber quién va ganando, es relevante. Unas elecciones pueden alterar el equilibrio en un país. En sistemas parlamentarios como el nuestro, además, puede ser un elemento para decidir tu voto. Si tienes varios candidatos puedes decidir votar a tu segunda opción si al otro le ves más fuerte en las encuestas. O en circunscripciones pequeñas, puede servir para descartar a partidos que no tienen posibilidades.

En segundo lugar, a veces se nos olvida, pero la gente lee el periódico para entretenerse. Estoy muy convencido de esto. Si el periódico fuera muy aburrido, la gente no lo querría leer. La mayoría de la gente lo lee en su tiempo libre. Entonces, yo creo que las elecciones tienen un poco de partido de fútbol. A la gente le gusta verlas en tiempo real, ver cómo cambian, tener expectativas de qué va a pasar, discutirlo en el café... Hay una parte de retransmisión que las elecciones permiten. Esto Nate Silver lo cuida muy bien. Es un tema que se actualiza cada día prácticamente, que va cambiando. Salta una polémica como los correos electrónicos del FBI y puedes ir a la web a ver si efectivamente ha cambiado o no, ver cómo cambia después de un debate... Hay una parte de retransmisión que es, sobre todo, interesante.

Y en tercer lugar, como decía, un buen tratamiento de los números a la gente le da cierta confianza y a la gente le gusta tener una información veraz.

### **¿La interactividad es un punto a favor del periodismo de datos o no tiene tanta importancia?**

Sí con justificación. A mucha gente le gusta poder profundizar. Esto es poder ver la metodología detallada, poder ver algunos números, que el porcentaje que das no sea

solo cuánto va a tener cada partido sino la distribución por territorios... Dar un poco más de información y que se pueda explorar gusta.

Sí que ha habido cierto retroceso con la interactividad. Hace cinco o diez años la apuesta del periodismo de datos era mucho más clara por la interactividad. La idea de que la gente se hiciera su propia historia, pudiera explorar los datos, llegar a sus propias conclusiones... y esto se mantiene en algunos casos pero se usa menos de lo que se usaba hace una década. Yo creo que es porque en el fondo la labor del periodista que es encontrar lo más importante y sintetizarlo sigue teniendo valor. La mayoría de las piezas de datos se han movido a jerarquizar la información. Por ejemplo, los titulares de las gráficas casi siempre apuntan al *finding*, a lo que tienes que ver. Si el titular es “España es favorita”, el titular te está diciendo que te fijes en esa información. Pero ellos ya te dicen lo que tienes que buscar, ya te hacen el resumen. Hubo un momento hace cinco o diez años en el que todo era periodismo exploratorio. Algo grande que te permitía ordenar 80 cosas. Sigue teniendo algunos usos pero es fácil caer en el abuso.

Quien lleva esto más a rajatabla es The New York Times. Hace pocos interactivos y los que hace son muy especiales. La mayoría de sus gráficos del día a día, aunque podrían hacerlos interactivos porque tienen los recursos, son estáticos. Sin ningún botón para tocar nada. Ellos pintan lo que quieren que veas.

Luego sí que hay casos particulares como el que está llevando a cabo FiveThirtyEight con el Mundial de fútbol que son un juego. Esta parte de que el lector pueda jugar a hacer su propia predicción. Pero al final se queda en eso, son casos concretos. Estoy seguro de que el porcentaje de gente que los acaba usando es bastante pequeño.

### **¿Cuáles son los mayores retos a los que se enfrentan los modelos predictivos?**

El principal reto ahora, que no lo sabíamos en 2016 pero lo sabemos desde entonces, es cómo comunicar las predicciones probabilísticas. En el momento en el que nos ponemos a hacer predicciones probabilísticas tipo “Donald Trump tiene un 30% de posibilidades de ganar” y eso luego ocurre, la gente se enfada. Esto no es razonable. No es razonable porque no hemos hecho el modelo para decir que va a ganar Hillary Clinton. Lo hemos hecho para mostrar que tiene un 70%. Eso quiere decir que 3 de cada 10 veces, que son muchas, ganará Trump. Esa no fue la interpretación de una gran parte de los lectores. Y, más que de los lectores, del mundillo de medios de Estados Unidos.

Es el cabreo gigante que tiene (Nate) Silver en 2016. Estamos diciendo que esto puede pasar, ha pasado y se está diciendo que no está bien cuando nosotros lo que podemos decir es que el modelo sí que está bien. Que esté bien o esté mal no se mira así, se mira con muchas más observaciones. El gran reto es ese y es la misma lectura que hizo The New York Times en las mismas elecciones. Fallamos en explicar que un 80% no es un 100%. Esto parece loco. ¿En serio hay que explicar que un 80 no es un 100? Por eso en 2016 pasa todo esto, la gente que manejaba este tipo de modelos no se esperaba este problema de comunicación. ¿Cómo arreglarlo? No lo sé, la gente está viendo cómo visualizar los datos, hay que ser bastante cautos. Yo soy bastante pesimista en que se puedan hacer las cosas muy distintas. En aquel momento ya usábamos símiles del tipo: “La probabilidad de que gane Trump es la misma que tiene Messi de fallar un penalti”. Todo el mundo sabe que Messi falla penaltis de vez en cuando. Lo veo complicado.

**En el modelo del Mundial que publicó en El País aparecía debajo una explicación que decía que el favoritismo de Brasil no significaba que fuera a ganar. Que un 17% es como tirar un dado y que salga el número que tú quieres.**

Exacto. Ahora, por ejemplo, eliminan a España y se dice que el modelo está mal. No, el modelo completo decía que había un 90% de que la eliminaran en algún momento del torneo. Es muy difícil porque es un sesgo humano. Confundimos ‘más probable’ con ‘muy probable’ y no es lo mismo. Que algo sea más probable que las alternativas no quiere decir que sea muy probable. ¿Cómo comunicar esto? Ya digo, creo que se va a hablar mucho de cómo hacerlo y en el fondo lo único que nos va a salvar es que la gente se familiarice con esto, que tenga experiencia. Esto se ve muy claro en deportes. La gente sí tiene esa percepción. Si el Real Madrid juega la final de Champions con un equipo inferior todo el mundo sabe que el Madrid es favorito pero nadie piensa que no pueda perder. ¿Por qué? Porque han visto muchos partidos y tienen la experiencia y piensan: “cuidado, que hemos sido favoritos otras veces y al final no hemos ganado”. La gente interpreta mejor los números donde tiene experiencia y eso en deportes se ve. En política hasta ahora la gente no tiene mucha pero a lo mejor Trump nos vale para que la próxima vez la gente diga “bueno, este es el favorito pero acuérdate de Trump”. Creo que la crítica ha sido excesiva pero a lo mejor es parte del proceso para que el público se familiarice con estos modelos.

**En política y deporte está asumido el uso de los datos. FiveThirtyEight tiene otras secciones como economía o cultura en las que en menor medida también los utiliza. ¿Qué aplicaciones pueden tener los datos en esas otras secciones que no son tan dadas a modelos predictivos o a mostrar la información de forma gráfica?**

Es algo que creo que va a ser más o menos transversal. Los datos van a ir entrando y han entrado ya en muchos temas. Política y deporte son los grandes nichos no tanto porque los datos sean particularmente útiles ahí sino porque son productos regulares y por lo tanto es más fácil invertir en desarrollo. Si tú tienes experiencia en hacer modelos de política, cada cuatro años hay elecciones generales y cada dos años las hay en otros lugares. Es una mejor inversión ponerte a hacer esos modelos que no un tema menos recurrente como puede ser la crisis migratoria que tenemos ahora. Más todavía con deporte porque las competiciones vuelven cada año. Yo tengo claro que los trabajos que tengo son en gran medida por eso. La predicción del Mundial a mí me cuesta mucho hacerla pero luego la puedo volver a usar. El conocimiento lo puedo usar. Para predecir las primarias del PP podríamos haber intentado algo, pero es algo puntual. ¿Datos en otros temas? En todo. Yo creo que la otra referencia (además de FiveThirtyEight), aunque no hace deportes, es The Upshot, la sección de datos de The New York Times. Por allí pasó Nate Silver y cuando se fue a ESPN se le dio un giro a la sección. No es solo sobre datos sino que también hay mucho periodismo analítico y se incluyen temas como debates sobre sanidad, inmigración o igualdad. Es mi referencia y yo creo que ahí van a haber datos para todo aquello que le puede interesar a un periódico.

**¿El impacto que tuvo el modelo de Nate Silver en 2008 es lo que ha generado que estos modelos probabilísticos se hayan vuelto tan populares y que se reproduzcan en otros países como en España con autores como usted?**

Sí. El trabajo de Silver entre 2008 y 2012 es la referencia. Había algunos trabajos académicos haciendo esto pero no había mucha tradición de hacer esto en la universidad ya que esta se encarga más de la explicación, no de la predicción. Entonces la referencia es seguramente Silver aunque hubo propuestas de otra gente. ¿Por qué en ese momento? Como te digo, yo creo que se debe a estas razones de oferta y demanda. Primero, esto se empieza a poder hacer porque hay muchas encuestas, se podía haber hecho antes pero de forma mucho más manual. Y luego se da que hay gente capaz de hacer esto. Es un



fenómeno ligado a internet. En un periódico tradicional era muy difícil que hubiera gente con un manejo de los datos que permitiera hacer esto. ¿Qué pasa entre 2000 y 2010? Irrumpe internet y los frikis de las estadísticas se ponen a colgar sus cosas o tienen más incentivos para hacerlo. Antes no ibas a invertir tu tiempo para predecir las elecciones y contárselo a la gente de un bar, pero ahora tienes la posibilidad de colgarlo en internet, que lo vea más gente y se monte una comunidad. De ahí viene Silver, de un blog de frikis de béisbol. La gente empieza a ver esto y le gusta. Ese es el camino que también sigo yo. Yo no soy periodista y he llegado a los periódicos a raíz de escribir en blogs. Hacía numeritos por ahí, empecé a contar cosas en 2006 con un blog sobre datos y actualidad con esta idea de que lo hacía porque de golpe se podía hacer. Silver es la referencia porque era Estados Unidos, esas elecciones tenían mucho interés y su modelo se ha ido replicando. Seguramente va a haber más. En España está yendo un poco lento pero esto se va a quedar.

### **¿Qué habilidades necesita un periodista tradicional para poder manejar los datos?**

Hay distintos niveles. El grueso de los periodistas de datos no programan. Se puede empezar manejando hojas de cálculo, viendo distintas herramientas, leyendo sobre el tema. Por decir una aproximación, 9 de cada 10 periodistas de datos en España es lo que manejan. Luego puede que tengan un desarrollador en el equipo. Puedes llegar muy lejos con eso. Si quieres apostar más fuerte, la clave que te abre la puerta a hacer más cosas es saber programar. Aprender un lenguaje de programación como R o Python. ¿Cómo aprender? Esto ahora es relativamente fácil. Fácil nunca es porque el primer lenguaje de programación es una cosa un poco marciana y siempre cuesta un poco empezar, pero hay un montón de recursos en internet. Una vez aprendes uno ya tienes la puerta abierta. Yo programaba en un lenguaje de programación desde hacía muchos años por mi trabajo en la universidad pero con R que es el que utilizo ahora empecé hace dos años. Ahora estoy aprendiendo otros para visualización que no voy a controlar porque son para especialistas, pero que quiero saber un poco. Entonces, ¿conviene saber programar? Es una inversión más grande pero es algo muy demandado. En España menos, pero las referencias como FiveThirtyEight o The Upshot y medios como The Economist o Financial Times ya programan o por lo menos lo intentan.

**En las predicciones deportivas se utiliza habitualmente el Elo que es un sistema de**

**puntuación que viene del ajedrez. ¿Por qué se utiliza y cómo se adapta en otros deportes?**

El Elo es relativamente sencillo de calcular. Tiene unas propiedades matemáticas bastante chulas y esas son las principales razones. Es útil para la predicción y es sencillo de utilizar, aunque no en todos los campos. En fútbol es bastante complicado y, por ejemplo, FiveThirtyEight no lo usa, utilizan una especie de Elo diferente que no llaman Elo. Uno de los problemas es el empate, que es más fácil de solventar, y otro problema son los goles. El Elo está pensado para predecir si alguien gana, pierde o empatara, pero en fútbol es relevante saber el resultado exacto. Por ejemplo, en Champions League, el partido de vuelta de una eliminatoria va a estar totalmente condicionado por el primero y no tiene sentido simularlo ignorando el primer partido. Yo creo que son las razones principales por las que FiveThirtyEight no utiliza el Elo en fútbol y utilizan un sistema parecido que no llaman así que consiste en un valor ofensivo y otro defensivo. La gracia del Elo es que te permite mantener una clasificación sin tener una liga regular. Si todos los equipos juegan siempre entre todos es más fácil regularlo, pero, ¿es mejor el Madrid que el PSG? No lo podemos saber seguro porque no se enfrentan habitualmente ni tienen los mismos rivales. Ahí entra en juego el Elo porque te pone en la pista. Los puntos de la clasificación Elo pueden viajar entre ligas y entre continentes sin que los equipos se enfrenten entre sí. Pero sigue siendo una dificultad grande.

**FiveThirtyEight utiliza los datos como nicho pero siempre ha estado ligado a un gran medio, ya sea The New York Times, ESPN o ABC News. ¿Usted cree que es viable un medio de comunicación de este estilo que funcione con independencia?**

Es el mismo problema que tienen todos los medios. Es una audiencia de nicho. Por ejemplo en España, que es un país pequeño o bastante más que Estados Unidos, tendrías el potencial de la audiencia en español pero producir periodismo de datos es caro y pagarlo con publicidad es inviable. Los productos de suscripción y de nicho ahora mismo son una cosa complicada. ¿En algún momento? Puede ser. Al final incluso FiveThirtyEight es propiedad de una gran compañía de televisión. Mismo caso que The Upshot en The New York Times aunque seguramente ellos se pagan a sí mismos porque son un producto de calidad y tienen unos lectores fieles. No conozco las cuentas de FiveThirtyEight pero quizás ellos no se pagan a sí mismos. Es un producto gratuito

que solo tiene publicidad. ¿Da esa publicidad para pagar lo que cuesta hacer FiveThirtyEight? No lo sé. Lo veo más como un producto de prestigio para una televisión. Ya digo que no conozco sus cuentas.

Por otra parte, hay todo un mundo que está por ver. Los deportes mueven mucho dinero y seguirlos en vivo tiene mucho interés. Un *site* para seguir eventos deportivos pueden ser un negocio. Con los e-sports ya existe. La gente paga para ver vídeos de partidas. ¿Con el deporte existirá? Seguramente sí, lo que no sé es si lo hará un medio o lo hará una tecnológica.

### **¿Qué diría que tiene FiveThirtyEight que lo hace diferente a otros productos similares?**

Tiene básicamente a Nate Silver, una persona muy intuitiva con los modelos y que los ha diseñado con una metodología que me gusta mucho. Hay otras metodologías pero yo creo que él acierta con la que usa. Además, tienen un buen tono y una plantilla de gente muy talentosa. El tratamiento de los datos es muy bueno, eso es lo mejor que tienen. Su debilidad más grande es que a lo mejor caen demasiado en el nicho. Hacen cosas que no interesan a la suficiente gente. Es algo que a veces perdemos de vista los periodistas. Que algo me encante a mí no significa que le vaya a gustar también a la gente. FiveThirtyEight a veces peca de friki, que es normal porque recoge a casi todos los frikis, pero cuando planteamos llevar algo que hacen ellos a nuestra audiencia a veces nos tenemos que parar a pensar si tiene sentido y no siempre lo tiene.

### **8.3. Entrevista a John Burn-Murdoch**

#### **¿Por qué cree que los datos se han vuelto tan relevantes en el periodismo actual?**

Creo que en gran parte se debe a la disponibilidad de los datos. Continuamente leemos que cada vez hay más información y más datos en el mundo, y creo que es verdad. Eso hace que sea mucho más fácil encontrar datos sobre una gran variedad de temas que después podemos utilizar para encontrar historias. Por otra parte, puede que simplemente se deba a que hay más periodistas de datos. Es difícil decir si primero llegó el periodismo de datos o si lo hicieron antes los periodistas de datos. Posiblemente lo hicieron al mismo tiempo y como resultado ahora hay más datos disponibles y más

personas capaces de trabajar con ellos. También el *software* y las herramientas actuales permiten analizar una mayor cantidad de información en menos tiempo, mientras que hace años gran parte del periodismo de datos que se lleva a cabo hoy en día lo habría realizado un equipo de investigación o habría requerido mucho más tiempo. Así que la disponibilidad y las herramientas dan lugar a que haya mayor variedad de temas que se pueden tratar con datos y a que podamos hacerlo más rápidamente.

### **¿Cree que se debe a la falta de confianza de la audiencia en el periodismo reciente?**

Personalmente, no. Creo que no es lo habitual -aunque este comentario no lo verían bien muchos periodistas y medios- que lo que la audiencia quiere tenga un impacto real en lo que hacen las redacciones. Es cierto que el periodismo de datos se ha vuelto algo popular, ahora tenemos premios de periodismo de datos y ese tipo de cosas, que probablemente lleven a las redacciones a contratar más periodistas de datos y a crear equipos de periodismo de datos. Puede que también influya el hecho de que el personal sea más versátil y eficiente, pero no creo que esto tenga que ver con la confianza. Creo que la gente que desconfía del periodismo lo hace ya sea de datos o de cualquier otro tipo. Hay una expresión en inglés que dice que hay varios tipos de mentiras: las mentiras, las malditas mentiras y las estadísticas (*lies, damned lies and statistics*). Esta frase usa para decir que las estadísticas se pueden utilizar para manipular cualquier historia. Como periodista de datos no lo apoyo, para mí las mismas personas que dicen que no confían en los medios son las que ven una historia contada con datos y dicen “ah, bueno, esto es una historia de periodismo de datos, me están mintiendo con lo que sea”. Es cierto que ha habido un crecimiento del periodismo de datos al mismo tiempo que ha habido una disminución de la confianza, pero no creo que estén necesariamente relacionados.

### **Antes explicaba que cada vez hay más periodistas de datos. ¿Cuáles son las habilidades que un periodista necesita para trabajar con datos?**

Una habilidad fundamental es que el periodista se sienta cómodo al trabajar con números, haciendo cálculos. En mi caso, matemáticas fue siempre la asignatura que mejor se me daba en el colegio. No creo que sea necesario hacer un curso de periodismo de datos, aunque yo sí lo he hecho, sino simplemente estar cómodo al trabajar con datos. En las redacciones en las que he estado, las de The Guardian y The Financial

Times, los reporteros que trabajan habitualmente con datos son gente que ha estudiado matemáticas a un nivel decente en el instituto. A muchos periodistas les asustan los datos, y no se sienten cómodos trabajando con hojas de cálculo o similares. Son un perfil que dejó de estudiar matemáticas en cuanto pudo. Así que cuando aparecen esos extraños casos de periodistas que no tienen problemas con las matemáticas, suelen acabar haciendo bastante periodismo de datos. En el caso del Financial Times, muchos de nuestros periodistas de datos trabajaron antes en bancos o en otros negocios como contables y ahora son reporteros habituales en el FT. Son gente que se siente muy cómoda cuando se sienta, abre una hoja de cálculo y la utiliza para analizar una historia. Es algo que no le pasa a la mayoría.

Por otra parte, creo que el crecimiento del periodismo digital y el periodismo en internet hace que, cada vez más, las redacciones estén contratando a gente que tiene cierta experiencia con las matemáticas, ya sea por haber estudiado esa carrera, informática o alguna ingeniería que le da esas habilidades con los números. Aunque muchas de estas personas fueron contratadas originalmente para desarrollar las webs de los medios, en muchos casos han ido pasando a formar parte del equipo que se dedica al periodismo de datos. En el Financial Times, por ejemplo, tenemos en Londres unos 5 desarrolladores en nuestro equipo de noticias interactivas, que además de crear mapas y páginas interactivas, se dedican también al análisis de datos, sencillamente porque tienen la capacidad y se sienten cómodos trabajando con lenguajes de programación y hojas de cálculo. Así que en definitiva esos dos son los perfiles fundamentales: el periodista que se siente cómodo trabajando con números y datos y las personas con unas capacidades analíticas que acaban formando parte del equipo que se dedica al periodismo de datos.

**Decía al inicio que cada vez hay más herramientas disponibles. ¿Cómo afecta esto al desarrollo del periodismo de datos?**

Creo que este es un factor muy importante, no solo que las herramientas están disponibles sino que además son gratis en la mayoría de los casos, incluso lenguajes de programación como Python o R. Que sea gratis hace que la gente pueda entrenarse de forma muy rápida. Cuando empecé en el Financial Times en 2013 no había programado nunca, pero en unos dos años aprendí XML, CSS, Javascript, algo de D3 y actualmente estoy haciendo muchas cosas con R. El hecho de que pueda aprenderlos, tanto en el

trabajo como por mi cuenta, sin que tengan ningún coste para mí es una gran ventaja. Otras personas que no sean periodistas de datos, si deciden que quieren aprender por ejemplo Python para el trabajo, pueden hacerlo. El bajo coste de las herramientas es una gran ventaja y ocurre lo mismo con la visualización de datos, ya sea con D3, R o una herramienta como Tableau, la cual usan muchos periodistas de datos para aprender sobre visualización. Creo que esos dos factores influyen mucho. Por un lado que es muy fácil para las redacciones aplicar estas herramientas por su bajo coste y por otro que es sencillo de aprender para los periodistas ya sea en la oficina o en casa.

**¿Es la interactividad un punto a favor del periodismo de datos o realmente no es tan utilizada?**

Posiblemente no se está utilizando lo suficiente, especialmente en el período de los últimos tres años más o menos. Creo que las redacciones cada vez más se preguntan si la interactividad está justificada en tanto a la experiencia añadida que los usuarios reciben. Hay que tener en cuenta que crear un elemento interactivo requiere de mucho tiempo y de comprobaciones tales como que funcione a la perfección en diferentes tipos de dispositivos. Todo ese trabajo extra que es necesario para montar un gráfico interactivo no siempre está justificado, y las redacciones están reservándolos para esas ocasiones en las que se realiza una pieza que es realmente especial. Por ejemplo, cada vez se están elaborando menos mapas que tengan una gran profundidad y la interactividad se está llevando más hacia la ludificación. Aún así, la interactividad sigue siendo importante. The New York Times, que sigue siendo el líder, sigue trabajando mucho con la interactividad, como por ejemplo con mapas para realizar un seguimiento de las elecciones políticas. Creo que la interactividad sigue siendo un ente que está creciendo y que seguirá formando parte del periodismo, no solo de datos, y las organizaciones están siendo más cautas y están justificando más cuándo la utilizan para obtener los mejores resultados.

**FiveThirtyEight es un medio de comunicación que solo elabora piezas basadas en datos, pero depende de una gran corporación como es ABC News. Como experto en la materia, ¿crees que es viable un medio que tenga los datos como nicho con la audiencia y la capacidad económica suficiente para ser independiente?**

Mi respuesta simple y directa es no. Muchas veces hemos tenido ese debate en el FT, no

en tanto a crear un medio independiente sino a desarrollar proyectos como una *newsletter* para los lectores que esté basada en gráficos y datos. Pero cuanto más lo pensamos y analizamos nos damos cuenta de que la gran mayoría de los lectores están interesados en un tema concreto, ya sea economía, política o medio ambiente; y ya sea en formato texto, *podcast* o gráfico lo que quiere el lector es informarse sobre esos temas. Normalmente un lector no nos pide que le enseñemos cualquier gráfico o cualquier mapa sino solo aquellos que están relacionados con el tema del que quieren saber más. Obviamente los que nos dedicamos a esto y estamos en el entorno del periodismo de datos sí que consumimos datos y gráficos en general, pero cuando quieres crear un proyecto que tenga la audiencia o los suscriptores suficientes tiene más sentido hacerlo en torno al tema que en torno al formato. Sí que hay proyectos que funcionan pero con formas de financiación muy innovadoras y equipos pequeños. FiveThirtyEight no creo que funcionase sin el apoyo de una gran compañía.

**¿Qué considera que hace FiveThirtyEight que sea diferente a otros medios que hacen uso del periodismo de datos?**

Solamente es mi opinión, y conozco solo a un par de personas de las que trabajan en FiveThirtyEight, así que desde la lejanía y solo desde mi punto de vista no creo que hagan nada que sea muy distinto a lo que realizan otras redacciones. No es una crítica, su trabajo es excelente, pero aquello en lo que más destaca como es el análisis político y los modelos de predicción también se ve ahora en otros medios como The New York Times o The Economist por dar dos ejemplos. Los modelos de FiveThirtyEight son realmente complejos y es un trabajo muy bueno, pero incluso en el Financial Times estamos empezando a utilizar técnicas similares. Además, FiveThirtyEight realiza muchas piezas que son reportajes estándar que podrían elaborarse sin datos. Insisto en que su trabajo es asombroso pero no creo que sea especialmente diferente al que se está haciendo en las secciones de datos de otros grandes medios.

#### **8.4. Entrevista a Daniele Grasso**

**¿Por qué el periodismo de datos tiene tanta relevancia actualmente?**

Bueno, los datos siempre tuvieron relevancia, no es que la tengan ahora, lo que pasa es que durante mucho tiempo se destacaron menos, ya que los medios tradicionales

tardaron en elaborar periódicos digitales que no fueran copias de lo que hacían en papel, sobre todo en países como España. Entonces, que toda esa cantidad de datos se empezara a visualizar y a tratar debidamente fue algo que llegó tarde, pero los datos en periodismo existen desde que existe el periodismo.

**¿Qué motiva a los medios a utilizarlos? ¿Qué ventaja tiene que la fuente de información sean los datos?**

Lo que motiva a los medios no lo sé, depende de cada caso. No es que un director se levante por la mañana y diga: ahora todos a usar datos. No, depende de cómo trabaja cada uno. Hay mucha gente que no es capaz ni le gusta. La ventaja que tiene sobre otra fuente es que hay una metodología detrás, con una recopilación de datos. No es lo mismo que hablar con una persona que te puede mentir o decirte una cosa y querer hacerte entender otra. Obviamente unos datos pueden estar equivocados y pueden mentirte también, pero, al menos, se puede ver cómo han sido recogidos y ser consciente de para qué se están usando.

**¿Crees que esto está relacionado con la falta de confianza que los consumidores tienen en los medios, según varias encuestas?**

Sí, es una de las ventajas que da, el hecho de que el lector pueda revisar la metodología y saber de dónde sale la información. Pero no es una solución, como estamos viendo. Solo un porcentaje de las personas leen esto. Que el periodista sea transparente y esté dando la metodología, en teoría, debería de ser un factor que ponga más en valor lo que ha hecho.

**¿Qué habilidades se necesitan para trabajar como periodista de datos?**

En general, tengo dudas con eso de que se pueda hablar del periodista de datos. Es verdad que hay cada vez es algo con más especialización, pero no se deja de ser periodista, a secas. No se necesita nada en concreto más allá de tener la mente abierta a usar herramientas de diferente tipo. La cuestión básica es ser buen periodista, luego viene todo lo demás. Las herramientas se pueden aprender, la manera de utilizar una metodología u otra se puede aprender... pero la base es la misma.

**¿Cómo afecta a los medios que muchas de las herramientas sean gratis o de bajo coste?**



A los medios les beneficia, ya que si una herramienta es gratuita es mucho más fácil utilizarla. Pero creo que sobre todo influye a las personas que se quieren formar por su cuenta, que sean autodidactas, ya que no se van a encontrar muchas dificultades más allá de lo que implica ponerse a estudiar y aprender cosas nuevas.

**¿Cómo se utiliza la interactividad actualmente? ¿Es un factor importante o ha perdido importancia?**

No creo que se pueda hablar de un patrón común, porque no hay ninguna organización que lo regule ni hay un acuerdo ahí. Hay medios que apuestan más por ella y otros que menos. Yo creo que ha habido una clave importante que es el auge del teléfono móvil. En el móvil tienes que hacer que la gente vea la información en un espacio muy pequeño y hay que adaptarse, por eso se ha renunciado en muchos casos a la interactividad. El móvil implica un tamaño que no es tan grande, una carga de red que no siempre se puede soportar... Nosotros (en El Confidencial) sí que estamos intentando volver a cuanto más estático mejor, pero es un proceso en curso, no es un punto de llegada ni mucho menos.

**FiveThirtyEight es un medio que solamente utiliza los datos como fuente pero depende de una gran corporación como ABC News. ¿Crees que es posible un medio como este que tenga la capacidad de ser independiente?**

Se puede hacer la misma pregunta sin poner lo de los datos. Depende de que haya buen contenido en el medio, da igual que esté hecho con datos o con lo que sea. El problema es si un medio con un buen contenido puede sobrevivir siendo totalmente independiente. Es la pregunta del millón. Si pudiera contestar y estuviera seguro de que sí, ya habría montado el mío con cuatro colegas. Cualquiera que conteste dando una respuesta segura sobre esto miente, porque hay ejemplos de una cosa y ejemplos de la otra. FiveThirtyEight ha demostrado algunas cosas. Para empezar, ha recordado a todo el mundo que en Estados Unidos hay dinero y en Europa no. Cuando algo se compara a Estados Unidos siempre hay que tener presente que se habla de otra escala de dinero. Sería como comparar, sin menospreciar, los medios de España con los de Kosovo. Diríamos que los medios españoles son mucho mejores, claro. En el departamento de infografía de The New York Times trabajan 45 personas, en España el más grande tiene 6. Estoy muy a favor de tener a estos medios como referencia absoluta, pero no

podemos comparar. FiveThirtyEight es un experimento muy interesante. También es cierto que ellos están muy especializados en deportes, pero sus deportes tienen mucho que ver con la estadística. El béisbol sin estadística no tiene sentido casi. Por mucho que la puedas aplicar al fútbol no es lo mismo, no es un deporte tan aritmético como el béisbol o el fútbol americano. Los intentos que se han hecho aquí no terminan de funcionar. Es un caso muy americano el de FiveThirtyEight.

### **¿Qué crees que lo que diferencia a FiveThirtyEight?**

Volcar todo, absolutamente todo, desde un punto de vista de ciencia de datos. Es un poco la clave. Su apuesta es aplicar los métodos científicos de datos para crear historias periodísticas. Yo consumo bastante periodismo de datos, y a mí FiveThirtyEight, salvo algunas cosas concretas, que son absolutamente espectaculares, me parece que no es un periódico por el que yo pagaría por leer su contenido. Ellos están muy especializados y han encontrado ese filón, esto es una opinión personal. La clave es que han empezado a aplicar a absolutamente todo la metodología de la ciencia de datos.

### **¿En qué lugares crees que se está innovando más dentro del periodismo de datos?**

Diría que ninguno en concreto. Todo el mundo está peleándose con el móvil y está aprendiendo a utilizar, cada vez más, bases de datos muy grandes. La tendencia es manejar bases de datos cada vez mayores para encontrar historias periodísticas. Se han hecho intentos de aplicarlos a la realidad virtual o similares, pero no han tenido éxito. Entendiendo por éxito que haya tenido lectores, que haya tenido un sentido económico, un impacto a nivel económico y social, no que haya sido bonito.

## **8.5. Entrevista a Alberto Cairo**

### **¿Por qué cree que el periodismo de datos tiene tanta relevancia? ¿Es una cuestión actual o siempre ha sido relevante?**

Por lo menos en Estados Unidos, el uso de datos en periodismo no es algo nuevo. Los datos están presentes en los medios de comunicación desde los años 60. Hay una tradición en Estados Unidos que se inició hace cuatro o cinco décadas que se llama 'periodismo de precisión'. Fue iniciada por Philip Meyer, que estudió estadística en Harvard, cuando se le ocurrió que era posible aplicar métodos de las ciencias sociales,

de análisis cuantitativo, al periodismo.

En paralelo, hubo otra tradición en Estados Unidos que se llama 'periodismo asistido por computadoras', que consiste en el uso de computadoras para analizar números. Unos análisis que hoy en día nos parecen súper simples, claro, porque la capacidad de los ordenadores de la época era muy limitada. En esta época el uso de los datos consistió en analizar el significado de los números para luego escribir historias. Los periodistas que hacían este tipo de trabajo no hacían gráficos, por ejemplo.

En los últimos diez o veinte años se ha dado una confluencia de factores que ha impulsado un nuevo uso de los números dentro del periodismo. Estos factores son, por una parte, que los ordenadores son más baratos y más fáciles de utilizar. Por otra, que el *software* se ha hecho mucho más potente y en muchos casos es gratuito y de código abierto como R. También hay muchos más datos disponibles en formato digital que nunca. Por último, se da que a la audiencia, al menos en Estados Unidos, le gustan los gráficos y las representaciones estadísticas.

Todo esto lleva a que muchos medios, como The New York Times, hayan creado secciones específicas para el periodismo de datos. El Washington Post también los trata mucho. Y luego está FiveThirtyEight, que se dedica exclusivamente a eso. También ProPublica, que para mí es el ejemplo perfecto.

Además, también hay que tener en cuenta la utilidad. Los datos son una fuente más de información. Antes, la mayor parte del periodismo se hacía a través de hablar con gente, de hablar con fuentes. Los datos son otra fuente a la que puedes interrogar y de la que puedes extraer historias.

**¿Cree que la importancia de los datos tiene relación con la falta de confianza que, según varias encuestas, los consumidores tienen en los medios?**

La verdad es que no lo sé. No creo que el hecho de usar más números vaya a conducir a que las personas tengan más confianza en los medios de comunicación. También es cierto que yo soy bastante escéptico con estas encuestas de confianza. Es cierto que la confianza promedio ha caído en los últimos años, pero no ha caído en todos los sectores de la población. Hay una diferencia educativa: la gente que no lee medios de comunicación es la gente que más desconfía de los medios de comunicación. Así que no creo que el periodismo de datos pueda hacer que se recupere esa confianza, no deja de

ser otro tipo de periodismo hecho por los mismos medios de comunicación. Bien es cierto que la gente que, en general, sabe poco de números, tiene una especie de confianza ciega en ellos, como si fueran magia, la realidad. Eso es un problema y hay que matar ese mito, pero la tendencia existe.

**Ha comentado al inicio que el hecho de que las herramientas sean más accesibles facilita el periodismo de datos. ¿Qué habilidades necesita un periodista para poder manejar estas herramientas?**

Las herramientas son muy fáciles de aprender. Se puede hacer periodismo de datos con Excel. No se necesita aprender sobre programación avanzada. Incluso los periodistas que utilizamos lenguajes de programación como R, los utilizamos de forma muy limitada. Para dedicarse al periodismo de datos no se requieren unas habilidades tecnológicas muy avanzadas, se puede aprender en un mes o dos de práctica con herramientas. La parte complicada, donde es muy fácil meter la pata, es la otra, la que consiste en razonar con los números. Para mí lo más importante no es aprender a usar herramientas sino la parte conceptual, la parte de pensamiento.

**En los últimos tres años, el uso de la interactividad ha caído considerablemente, ¿a qué cree que se debe?**

Se debe a los móviles, básicamente. Hasta hace diez años, la gente cuando veía un medio de comunicación en internet lo veía en la pantalla del ordenador. Pantalla grande, cursor, ratón... puedes interactuar. Luego llegan los móviles, ¿qué pasa? Que las posibilidades de interactuar se limitan mucho porque las pantallas son más pequeñas. Al mismo tiempo, los medios de comunicación se dieron cuenta de que la mayoría usuarios no interactúan nada, simplemente abren la noticia, la leen, miran las imágenes y los gráficos y toda la interactividad que realizan es la de subir y bajar. Esto sobre todo lo comunicó The New York Times, que dijo que el 80% de sus usuarios eran así. Entonces no merece la pena dedicar una cantidad de tiempo y esfuerzo enorme en utilizar tanta interactividad para que luego los usuarios no la utilicen. Ahora bien, se ha mantenido en casos concretos. Por ejemplo, en las elecciones norteamericanas a los lectores no les interesa sólo el panorama nacional, les interesa llegar hasta los resultados de su propio barrio, y ahí sí ponen interactividad. No me parece malo que haya caído la interactividad, es como cualquier otra herramienta, hay que utilizarla cuando tenga

sentido.

**¿En qué apartados cree que se está innovando más dentro del periodismo de datos?**

Mi especialidad es la parte gráfica, así que no puedo hablar del periodismo de datos en general. Sé que en otras áreas hay mucha discusión en el uso de *machine learning*, por ejemplo, que básicamente son técnicas de inteligencia artificial, pero realmente no sé qué están haciendo, más allá de desarrollar algoritmos. En la parte gráfica, hay muchos experimentos ahora con realidad virtual y realidad aumentada, aunque no sé a dónde conducirá. Hay mucho trabajo también con los métodos predictivos, pero no sé realmente hasta qué punto se sigue innovando, la mayoría del trabajo es el que se venía haciendo con ponderación de encuestas y estadística bayesiana, creo que no hay demasiada innovación.

Desde un punto de vista personal, la innovación en periodismo no debería centrarse en crear herramientas o métodos nuevos, sino en centrarse en entender los que tenemos. A día de hoy sigo viendo muchas barbaridades en el uso de mapas y gráficos, por lo que lo importante no es usar nuevos gráficos, sino usar los que tenemos bien. Creo que pasa lo mismo en la parte predictiva, la gente dice que fallaron los modelos, pero no, los modelos estaban bien, pasó algo (con la victoria de Donald Trump en las elecciones estadounidenses de 2016) que tenía una probabilidad baja, pero no cero. Seguramente el esfuerzo en innovación que hay que hacer actualmente es explicar al gran público los resultados de estos modelos.

**Comparte la opinión de Kiko Llaneras, al que entrevisté y decía lo mismo acerca de la innovación en los modelos predictivos, que falta saber explicar los datos.**

Claro, hay autores que se dedican a explicar cómo comunicar riesgo. Los periodistas no leemos a esos autores y hay que leerlos más. La última vez que hablé con Kiko Llaneras le hablaba de esto. Yo no soy un experto en modelos y no puedo hablar de su calidad, pero sí lo soy en comunicación. Lo que dicen autores como Gerd Gigerenzer, experto en comunicar riesgo, es que no hay que usar porcentajes con decimales. No digas que algo tiene una probabilidad del 25,8%, porque da la impresión de ser muy preciso y de que es la verdad. Ese número que te sale en R lo tienes que ponderar. Algo muy sencillo es cargarse los decimales: el 25%. Ahí ya das la impresión de que la precisión es efecto

del cálculo. Y todavía más, no usar los porcentajes. Decir: uno entre cuatro. O, por ejemplo, para un 15%, decir que es lo mismo que lanzar un dado y que salga la cara que quieres. Unas elecciones hay que plantearlas en estos términos, como lanzar un dado, porque la realidad es aleatoria. “Hay un 30% de posibilidades de que Donald Trump gane”, vale, pero se puede decir que Trump tiene tantas posibilidades como tirar un dado y que salga el 1 o el 2. De esa manera yo creo que se entiende mejor, porque un 30% es una probabilidad grande. Yo veo que los periodistas de datos siguen utilizando mucho lo decimales, que no está mal si sabes interpretar los números, pero no todo el mundo lo sabe realizar.

ProPublica y FiveThirtyEight hacen esto muy bien. ProPublica hace unas historias muy complejas pero te explica lo que hay detrás. FiveThirtyEight si en una noticia habla del margen de error te explica lo que es el margen de error, no te dicen sólo que es de 3 puntos, y te lo explican con un lenguaje comprensible, además. Es diferenciarse del periodismo tradicional, que asume que el lector es idiota y no puede aprender, por lo que no se le explica nada, solo se muestra lo que el periodista cree que es importante. El periodismo moderno tiene que asumir que el lector no es idiota y que puede aprender. ¿Cómo? Si le explicamos. Si tienes que hablar de la incertidumbre en los datos, explica qué es y hazlo parte de la noticia. Estoy siendo injusto porque hay periodistas tradicionales muy buenos y modernos muy malos, pero hablo en general.

### **¿Qué importancia tiene utilizar los datos a la hora de analizar a la audiencia?**

Sobre esto, creo que se cae mucho en el error de utilizar las analíticas para buscar a tu tipo de lector. Y no, no hay un tipo de lector, hay varios tipos de lectores. A mí, por muy embellecida y llena de gráficos que me pongan una noticia sobre deportes no me va a interesar, en cambio si es de política o análisis sí. ¿Significa esto que el lector de deportes es peor que yo? No, significa que hay varios tipos de lectores. Hay medios que esto lo tienen más fácil porque la audiencia ya la tienen segmentada y, entonces, lo hacen muy bien como FiveThirtyEight. Ellos saben que su audiencia está basada en gente friki y hacen el contenido orientado a ellos. También ProPublica, es un medio cuyos lectores suelen tener un alto nivel de estudios y están interesados en política. La dificultad y donde no se está haciendo bien es en los medios generalistas.

### **¿Es la audiencia la que permite a FiveThirtyEight basarse en los datos y no en**

### **distintos temas como política o deporte?**

Sí, al final se trata de la audiencia, y la de FiveThirtyEight es diferente. Los lectores de este medio son lo que en Estados Unidos se conoce como *wonks*, los frikis. Al que le gusta la sección de deportes de FiveThirtyEight seguramente le gustan tanto los números del deporte como el deporte en sí. Le gusta ver todas las tablas de números que tiene el béisbol, por ejemplo. Son audiencias diferentes.

### **¿Qué diferencia a FiveThirtyEight de otros medios que también trabajan con periodismo de datos?**

Varias cosas. Por un lado, la audiencia segmentada, que tiene un interés muy específico y le gustan los números y el análisis. Y luego, el hecho de que desde el principio la gente que trabajaba en FiveThirtyEight era especialista en datos y estadística. Los medios tradicionales para hacer este tipo de contenido tienen que contratar a gente nueva, pero la gente que contrata a otra en FiveThirtyEight es gente que ya trabaja con este periodismo, como Nate Silver. En FiveThirtyEight, Nate Silver es el jefe, y los subdirectores seguramente también son frikis de los números. Incluso los pocos medios tradicionales que han ascendido a especialistas en datos a puestos de poder, como The New York Times y el Washington Post, están en desventaja.

### **¿Cree que sería posible desarrollar un medio de comunicación con los datos como nicho que fuera viable económicamente? FiveThirtyEight, por ejemplo, siempre ha dependido de una gran compañía.**

Realmente no lo sé. En Estados Unidos hay distintos modelos de negocio. FiveThirtyEight, como dices, sigue bajo el paraguas de ESPN y ABC. Lo que sí hay en Estados Unidos son medios de comunicación sin ánimo de lucro, como el Texas Tribune o ProPublica. Son medios grandes y potentes que continúan creciendo y contratando y en ambos el periodismo de datos es central. Según datos del propio Texas Tribune, el 60% de su tráfico viene de bases de datos. Por anuncios lo veo difícil, yo creo que ya no sería posible, aunque FiveThirtyEight creo que sólo depende de ellos.